

# ASYMPTOTICS OF FLUCTUATIONS IN CRUMP–MODE–JAGERS PROCESSES: THE LATTICE CASE

By SVANTE JANSON

---

## Abstract

Consider a supercritical Crump–Mode–Jagers process in which all births are at integer times (the lattice case). Let  $\hat{\mu}(z)$  be the generating function of the intensity of the offspring process, and consider the complex roots of  $\hat{\mu}(z) = 1$ . The root of smallest absolute value is  $e^{-\alpha} = 1/m$ , where  $\alpha > 0$  is the Malthusian parameter; let  $\gamma_*$  be the root of second smallest absolute value. Subject to some technical conditions, the second-order fluctuations of the age distribution exhibit one of three types of behaviour: (i) when  $\gamma_* > e^{-\alpha/2} = m^{-1/2}$ , they are asymptotically normal; (ii) when  $\gamma_* = e^{-\alpha/2}$ , they are still asymptotically normal, but with a larger variance; and (iii) when  $\gamma_* < e^{-\alpha/2}$ , the fluctuations are in general oscillatory and (degenerate cases excluded) do not converge in distribution. This trichotomy is similar to what has been observed in related situations, such as some other branching processes and for Pólya urns. The results lead to a symbolic calculus describing the limits. The asymptotic results also apply to the total of other (random) characteristics of the population.

*Keywords:* Crump–Mode–Jagers process; age distribution

2010 Mathematics Subject Classification: Primary 60J80

Secondary 60F05

---

## 1. Introduction

Consider a Crump–Mode–Jagers branching process, starting with a single individual born at time 0, where an individual has  $N \leq \infty$  children born at the times when the parent is of age  $\xi_1 \leq \xi_2 \leq \dots$ . Here  $N$  and  $(\xi_i)_i$  are random, and different individuals have independent copies of these random variables. Technically, it is convenient to regard  $\{\xi_i\}_1^N$  as a point process  $\Xi$  on  $[0, \infty)$ , and give each individual  $x$  an independent copy  $\Xi_x$  of  $\Xi$ . For further details, see, e.g. [7].

We consider the supercritical case, for which, with positive probability, the population grows to infinity. As is well known, under weak assumptions, the population grows exponentially, like  $e^{\alpha t}$  for some constant  $\alpha > 0$  called the *Malthusian parameter* (see, e.g. [7, Theorems (6.3.3) and (6.8.1)]); in particular, the population size properly normalized converges to some positive random variable, and the age distribution stabilizes. Our purpose is to study the second-order fluctuations of the age distribution, or, more generally, of the population counts of a random characteristic.

In this paper we consider the lattice case; thus, we assume that the  $\xi_i$  are integer valued, so that all births occur at integer times a.s., but (without loss of generality) there is no  $d > 1$  such that all birth times a.s. are divisible by  $d$ .

Our setting can, for example, be considered as a model for the (female) population of some animal that is fertile several years and bears one or several children once every year, with the numbers of children in different years random and dependent.

Our main results show that there are three different cases depending on properties of the intensity measure  $\mathbb{E} \Xi$  of the offspring process. In Theorem 2.1 the fluctuations, properly normalized, are asymptotically normal, with only short-range dependence between different times. In Theorem 2.3 there is long-range dependence and, again after proper (but different) normalization, the fluctuations are a.s. approximated by oscillating (almost periodic) random functions of  $\log n$ ; these functions are essentially determined by the initial phase of the branching process, and are presumably nonnormal. Theorem 2.2 is an intermediate boundary case. Details are in Section 2.

A similar trichotomy has been found in various related situations. Similar results are proved for multitype Markov branching processes by Asmussen and Hering [1, Section VIII.3]. Their type space can be very general, so this setting also includes the single-type non-Markov case studied here (also in the nonlattice case [1, Section VIII.12]), since a Crump–Mode–Jagers branching process can be viewed as a Markov process in which the type of an individual is its entire life history up to the present. However, in general, this is a large type space, and the assumptions of [1] may not be satisfied; in particular, their ‘condition (M)’ [1, p. 156] is typically not satisfied, by the same argument as in [1, p. 173] for a related situation. Hence, we cannot obtain our results directly from the closely related results in [1], although there is an overlap in some special cases (for example, the Galton–Watson case in Example 2.1 below).

Another related situation is given by multicolour Pólya urn processes for which in [9] we used methods and results from branching process theory. The same trichotomy appears there too, with a criterion formulated in terms of eigenvalues of a matrix that can be seen as the ‘expected offspring matrix’ in that setting.

It would be interesting to find more general theorems that would present these different but obviously related results together in a common structure.

**Remark 1.1.** Our setup includes the Galton–Watson case, in which all births occur when the mother is of age 1 (Example 2.1), but this case is much simpler than the general case and can be treated by simpler methods; see [7, Section 2.10], where results closely related to those below are given.

**Remark 1.2.** It would be very interesting to extend the results to the more complicated nonlattice case; under suitable assumptions, we would expect similar results, but this case presents other technical challenges, and we leave it as an open problem.

## 2. Assumptions and main result

Let  $\mu := \mathbb{E} \Xi$  be the intensity measure of the offspring process; thus,  $\mu := \sum_{k=0}^{\infty} \mu_k \delta_k$ , where  $\mu_k$  is the expected number of children that an individual bears at age  $k$  (and  $\delta_k$  is the Dirac delta, i.e. a point mass at  $k$ ). Let  $N_k := \Xi\{k\}$  be the number of children born to an individual of age  $k$ . Then  $N = \sum_{k=1}^{\infty} N_k$  and  $\mu_k = \mathbb{E} N_k$ .

We make the following standing assumptions, valid throughout the paper. The first assumption (supercriticality) is essential, for otherwise there is no asymptotic behaviour to analyse. Assumptions (A2)–(A4) are simplifying and convenient, but presumably not essential. (For (A4), this is shown in Example 11.1.)

(A1) The process is supercritical, i.e.  $\mu([0, \infty]) = \sum_{k=0}^{\infty} \mu_k = \mathbb{E} N > 1$ .

(A2) No children are born instantaneously, i.e.  $\mu_0 = 0$ .

(A3)  $N \geq 1$  a.s. Thus, the process a.s. survives.

(A4) There are no deaths.

For all complex  $z$  such that either  $z \geq 0$  or the sums or expectations below converge absolutely, define

$$\hat{\mu}(z) := \sum_{k=0}^{\infty} \mu_k z^k = \sum_{k=0}^{\infty} \mathbb{E}[N_k] z^k = \mathbb{E} \sum_{i=1}^N z^{\xi_i}, \tag{2.1}$$

$$\hat{\Xi}(z) := \int_0^{\infty} z^x d\Xi(x) = \sum_{i=1}^N z^{\xi_i} = \sum_{k=0}^{\infty} N_k z^k, \tag{2.2}$$

so the complex-valued random variable in (2.2) has expectation  $\mathbb{E} \hat{\Xi}(z) = \hat{\mu}(z)$ .

We make two other standing assumptions.

(A5)  $\hat{\mu}(m^{-1}) = 1$  for some  $m > 1$ .

Thus,  $\alpha := \log m$  satisfies  $\sum_{k=1}^{\infty} \mu_k e^{-k\alpha} = \hat{\mu}(e^{-\alpha}) = 1$ ,  $\alpha$  is the Malthusian parameter, and the population grows roughly by a factor  $e^\alpha = m$  for each generation (see, e.g. (2.7) and (2.8) below).

(A6)  $\mathbb{E}[\hat{\Xi}(r)^2] < \infty$  for some  $r > m^{-1/2}$ .

In the sequel we fix some  $r > m^{-1/2}$  satisfying (A6), and assume for convenience that  $r \leq 1$ . It is crucial for our arguments that

$$\hat{\mu}(r) = \mathbb{E} \hat{\Xi}(r) < \infty, \tag{2.3}$$

a fact that we exploit immediately. Define  $D_r = \{z : |z| < r\}$ . Then, for  $z \in D_r$ , (2.3) implies that  $\hat{\mu}(z)$  and  $\hat{\Xi}(z)$  are well defined and analytic for such  $z$ . Since  $\hat{\mu}(z)$  is a strictly increasing function on  $[0, \infty)$ ,  $m^{-1}$  in (A5) is the unique positive root of  $\hat{\mu}(z) = 1$ . However,  $\hat{\mu}(z) = 1$  may have other complex roots; we show in this paper that the asymptotic behaviour of the fluctuations depends crucially on the location of these roots. Define

$$\Gamma := \{z \in D_r : \hat{\mu}(z) = 1\} \quad \text{and} \quad \Gamma_* := \Gamma \setminus \{m^{-1}\}; \tag{2.4}$$

$$\gamma_* := \begin{cases} \inf\{|z| : z \in \Gamma_*\} & \text{if } \Gamma_* \text{ is nonempty,} \\ \infty & \text{otherwise;} \end{cases} \tag{2.5}$$

$$\Gamma_{**} := \{z \in \Gamma_* : |z| = \gamma_*\}. \tag{2.6}$$

These sets may depend on the choice of  $r$ , but, for our purposes, this does not matter; recall that we assume that  $r > m^{-1/2}$ . Since  $\hat{\mu}(z)$  is analytic on  $D_r$ ,  $\Gamma$  is discrete and so, for  $\gamma_* < \infty$ ,  $\Gamma_{**}$  is a finite nonempty set  $\{\gamma_1, \dots, \gamma_q\}$ , say.  $\gamma_*$  is critical below.

Let  $Z_n$  be the total number of individuals at time  $n$ ; then, by (A2),  $Z_n$  equals the number of individuals born up to time  $n$ . By assumption,  $Z_0 = 1$ , and by setting  $Z_n := 0$  for  $n < 0$  we define  $Z_n$  for all integers  $n$ . It is well known that  $Z_n$  grows asymptotically like  $m^n$  as  $n \rightarrow \infty$  as, e.g. in [7, Theorem (6.3.3)],

$$\mathbb{E} Z_n \sim c_1 m^n \quad \text{for some } c_1 > 0. \tag{2.7}$$

Moreover, if  $\mathbb{E}[\hat{\Xi}(m^{-1}) \log \hat{\Xi}(m^{-1})] < \infty$ , and, in particular, if  $\mathbb{E}[\hat{\Xi}(m^{-1})^2] < \infty$ , which follows from assumption (A6), then, as  $n \rightarrow \infty$ ,

$$\frac{Z_n}{m^n} \xrightarrow{\text{a.s.}} \mathcal{Z} \tag{2.8}$$

for some random variable  $\mathcal{Z} > 0$  (see, e.g. [10]). In particular, it follows that, for any fixed  $k \geq 1$ ,

$$\frac{Z_{n-k}}{Z_n} \xrightarrow{\text{a.s.}} m^{-k} \quad \text{as } n \rightarrow \infty. \tag{2.9}$$

By (A4), the number of individuals of age greater than or equal to  $k$  at time  $n$  equals  $Z_{n-k}$ . From (2.9), we expect that, for large  $n$ , this should be about  $m^{-k}Z_n$ , so to study the fluctuations we define

$$X_{n,k} := Z_{n-k} - m^{-k}Z_n, \quad k = 0, 1, \dots \tag{2.10}$$

Note that  $X_{n,0} = 0$ .

We state our main results as three theorems, treating each of the cases  $\gamma_* > m^{-1/2}$ ,  $\gamma_* = m^{-1/2}$ , and  $\gamma_* < m^{-1/2}$  separately. In particular, note that Theorems 2.1–2.2 yield asymptotic normality of  $X_{n,k}$  when  $\gamma_* \geq m^{-1/2}$ . Proofs are given in later sections. We extend the results to random characteristics in Section 11.

From assumption (A6) and (2.2),  $\mathbb{E}N_k^2 < \infty$  for every  $k \geq 1$ . For  $j, k \geq 1$ , define

$$\sigma_{jk} := \text{cov}(N_j, N_k), \tag{2.11}$$

and, for  $z \in D_r$ ,

$$\Sigma(z) := \sum_{i,j} \sigma_{ij} z^i \bar{z}^j = \text{cov}\left(\sum_i N_i z^i, \sum_j N_j \bar{z}^j\right) = \mathbb{E}[|\hat{\Xi}(z) - \hat{\mu}(z)|^2]. \tag{2.12}$$

For  $R > 0$ , let  $\ell_R^2$  be the Hilbert space of infinite vectors

$$\ell_R^2 := \left\{ (a_k)_{k=0}^\infty : \|(a_k)_{k=0}^\infty\|_{\ell_R^2}^2 := \sum_{k=0}^\infty R^{2k} |a_k|^2 < \infty \right\}. \tag{2.13}$$

To simplify the notation, we often denote a vector in  $\ell_R^2$  by  $(a_k)_k$ .

We begin with the case  $\gamma_* > m^{-1/2}$ , which, by (2.4)–(2.5), is equivalent to

(B)  $\hat{\mu}(z) \neq 1$  for all complex  $|z| \leq m^{-1/2}$  except possibly  $z = m^{-1}$ .

**Theorem 2.1.** *Assume that (A1)–(A6) and (B) hold, i.e.  $\gamma_* > m^{-1/2}$ . Then, as  $n \rightarrow \infty$ , jointly for all  $k \geq 0$ ,*

$$\frac{X_{n,k}}{\sqrt{Z_n}} \xrightarrow{D} \zeta_k \tag{2.14}$$

for some jointly normal random variables  $\zeta_k$  with mean  $\mathbb{E} \zeta_k = 0$  and covariance matrix given, for any finite sequence  $a_0, \dots, a_K$  of real numbers, by

$$\text{var}\left(\sum_k a_k \zeta_k\right) = \frac{m-1}{m} \oint_{|z|=m^{-1/2}} \frac{|\sum_k a_k z^k - \sum_k a_k m^{-k}|^2}{|1-z|^2 |1-\hat{\mu}(z)|^2} \Sigma(z) \frac{|dz|}{2\pi m^{-1/2}}. \tag{2.15}$$

Convergence (2.14) holds also in the stronger sense that, for any  $R < m^{1/2}$ ,

$$(Z_n^{-1/2} X_{n,k})_k \xrightarrow{D} (\zeta_k)_k$$

in the Hilbert space  $\ell_R^2$ .

The limit variables  $\zeta_k$  are nondegenerate unless  $\Xi$  is deterministic, i.e.  $N_k = \mu_k$  a.s. for each  $k \geq 0$ .

Recall that joint convergence of an infinite number of variables means joint convergence of any finite set. (This is convergence in the product space  $\mathbb{R}^\infty$ , see [2].) Note that trivially  $\zeta_0 = 0$  (included for completeness).

The variance formula (2.15) can be interpreted as a stochastic calculus, where the limit variables are seen as stochastic integrals (in a general sense) of certain functions on the circle  $|z| = m^{-1/2}$ ; these functions thus represent the random variables  $\zeta_k$ , and, therefore,

asymptotically  $X_{n,k}$ ; moreover, they can be used for convenient calculations. See Section 10 for details.

We give two proofs of Theorem 2.1. The first, in Sections 4–5, is based on some approximations and the elementary central limit theorem for sums of independent variables; this proof is extended to random characteristics in Section 11. The second proof, in Sections 6–7, is based on a martingale central limit theorem, and is easily adapted to prove Theorem 2.2 below in Section 8.

Consider next the cases  $\gamma_* \leq m^{1/2}$ . Then  $\Gamma_{**} = \{\gamma_1, \dots, \gamma_q\}$  is a nonempty finite set. For simplicity, assume that the condition

$$\hat{\mu}'(\gamma) \neq 0, \quad \gamma \in \Gamma_{**}, \tag{2.16}$$

holds, so that the points in  $\Gamma_{**}$  are simple roots of  $\hat{\mu}(z) = 1$ ; the modifications in the case of multiple roots are left to the reader. (See Remark 3.3, and note the related results for Pólya urns in [9, Theorems 3.23–3.24] and [11, Theorems 3.5–3.6].)

**Theorem 2.2.** *Assume that (A1)–(A6) hold,  $\gamma_* = m^{-1/2}$ , and that (2.16) holds. Then, as  $n \rightarrow \infty$ , jointly for all  $k \geq 0$ ,*

$$\frac{X_{n,k}}{\sqrt{nZ_n}} \xrightarrow{D} \zeta_k \tag{2.17}$$

for some jointly normal random variables  $\zeta_k$  with mean  $\mathbb{E} \zeta_k = 0$  and covariance matrix given, for any finite sequence  $a_0, \dots, a_K$  of real numbers, by

$$\text{var} \left( \sum_k a_k \zeta_k \right) = (m - 1) \sum_{p=1}^q \frac{|\sum_k a_k \gamma_p^k - \sum_k a_k m^{-k}|^2}{|1 - \gamma_p|^2 |\hat{\mu}'(\gamma_p)|^2} \Sigma(\gamma_p). \tag{2.18}$$

For any  $R < m^{1/2}$ , convergence (2.14) also holds in the Hilbert space  $\ell_R^2$ .

The limit variables  $\zeta_k$  are nondegenerate unless  $\hat{\Xi}(\gamma_p)$  is deterministic for each  $\gamma_p \in \Gamma_{**}$ .

**Theorem 2.3.** *Assume that (A1)–(A6) hold,  $\gamma_* < m^{-1/2}$ , and that (2.16) holds. Then there exist complex-valued random variables  $U_1, \dots, U_q$  and linearly independent vectors  $\vec{u}_i := (\gamma_i^k - m^{-k})_k$ ,  $i = 1, \dots, q$ , such that, for any  $R < m^{1/2}$ ,*

$$\gamma_*^n \vec{X}_n - \sum_{i=1}^q \left( \frac{\bar{\gamma}_i}{|\gamma_i|} \right)^n U_i \vec{u}_i \rightarrow 0 \quad \text{a.s. and in } L^2(\ell_R^2). \tag{2.19}$$

Furthermore,  $\mathbb{E} U_i = 0$ , and  $U_i$  is nondegenerate unless  $\hat{\Xi}(\gamma_i)$  is degenerate.

Theorems 2.1–2.3 exhibit several differences between the three cases  $\gamma_* \leq m^{-1/2}$ ; (cf. similar results for Pólya urns in e.g. [9, Theorems 3.22–3.24]).

- (a) For fixed  $k$ , the fluctuations  $X_{n,k}$  are asymptotically normal when  $\gamma_* \geq m^{-1/2}$ , but (presumably) not when  $\gamma_* < m^{-1/2}$ .
- (b) For fixed  $k$ , the fluctuations are typically of order  $Z_n^{1/2} \asymp m^{n/2}$  when  $\gamma_* > m^{-1/2}$ , slightly larger (by a power of  $n$ ) when  $\gamma_* = m^{-1/2}$ , and of much larger order  $\gamma_*^{-n}$  when  $\gamma_* < m^{-1/2}$ .
- (c) When  $\gamma_* < m^{-1/2}$ , the fluctuations exhibit oscillations that are periodic or almost periodic in  $\log n$  (see [3]). (Note that  $\gamma_i/|\gamma_i| \neq 1$  in (2.19), since  $m^{-1}$  is the only positive root in  $\Gamma$ .)

- (d) When  $\gamma_* < m^{-1/2}$ , the almost-sure approximation result (2.19) implies both long-range dependence as  $n \rightarrow \infty$ , and that the asymptotic behaviour is essentially determined by what happens in the first few generations. In contrast, the limits in (2.14) and (2.17) are mixing (see the proofs), i.e. the results also hold conditioned on the life histories of the first  $M$  individuals for any fixed  $M$ , and thus also conditioned on  $Z_1, \dots, Z_K$  for any fixed  $K$ ; hence, when  $\gamma_* \geq m^{-1/2}$ , the initial behaviour is eventually forgotten. Moreover, for  $\gamma_* > m^{-1/2}$ , there is only short-range dependence (see Example 10.1), while the case  $\gamma_* = m^{-1/2}$  shows an intermediate ‘medium-range’ dependence (see Subsection 10.2).
- (e) When  $\gamma_* > m^{-1/2}$ , the limit random variables  $\zeta_k$  in (2.14) are linearly independent, as a consequence of (2.15). When  $\gamma_* \leq m^{-1/2}$ , the limits in (2.17), or the components of the sum in (2.19), span a (typically)  $q$ -dimensional space of random variables, and any  $q + 1$  of them are linearly dependent; see also Section 10.

**Remark 2.1.** Above we consider  $X_{n,k}$  for  $k \geq 0$ , i.e. the age distribution of the population at time  $n$ . We can also define  $X_{n,k}$  by (2.10) for  $k < 0$ : this means looking into the future and can be interpreted as predicting the future population. As shown in Section 10, (2.14)–(2.15) and (2.17)–(2.18) extend to all  $k \in \mathbb{Z}$  (still jointly), and, similarly, taking the  $k$ th component in (2.19) yields a result that extends to all  $k \in \mathbb{Z}$ .

This enables us, for example, to obtain (by standard linear algebra) the best linear predictor of  $Z_{n+1}$  based on the observed  $Z_n, \dots, Z_{n-K}$  for any fixed  $K$ .

**Example 2.1.** (*Galton–Watson.*) The simplest example is a Galton–Watson process, where all children are born in a single litter to a parent of age 1, so  $N_k = 0$  for  $k \geq 2$ . (But all individuals live forever in our setting. In the traditional setting, only the newly born are counted, i.e.  $Z_n - Z_{n-1}$ ; the results are easily transferred to this version.) Then  $N = N_1$ ,  $m = \mu_1$ , and  $\hat{\mu}(z) = mz$ . Hence,  $\Gamma = \{m^{-1}\}$ ,  $\Gamma_* = \emptyset$ , and  $\gamma_* = \infty > m^{-1/2}$ . We assume that  $\mathbb{E}N^2 < \infty$ ; then (A6) holds for any  $r$ ; we also assume that  $N \geq 1$  a.s. and  $\mathbb{P}(N > 1) > 0$ ; then (A1)–(A6) and (B) hold. Thus, Theorem 2.1 applies. We obtain, for example, with  $\sigma^2 := \text{var}(N) = \sigma_{11}$ ,

$$\begin{aligned} \text{var}(\zeta_1) &= \frac{m-1}{m} \oint_{|z|=m^{-1/2}} \frac{|z-m^{-1}|^2}{|1-z|^2|1-mz|^2} \sigma^2 |z|^2 \frac{|dz|}{2\pi m^{-1/2}} \\ &= \sigma^2 \frac{m-1}{m^4} \oint_{|z|=m^{-1/2}} \frac{1}{|1-z|^2} \frac{|dz|}{2\pi m^{-1/2}} \\ &= \frac{\sigma^2}{m^3}. \end{aligned}$$

This can be shown directly in a much simpler way; see [7, Theorem (2.10.1)], which is essentially equivalent to our Theorem 2.1 in the Galton–Watson case (but without our assumption (A3)).

**Example 2.2.** Suppose that all children are born when the mother is of age 1 or 2, so  $N_k = 0$  for  $k > 2$ . Then  $\hat{\mu}(z) = \mu_1 z + \mu_2 z^2$ , where, by assumption,  $\mu_1 + \mu_2 > 1$  and  $\mu_1 > 0$ . (A5) yields  $m^2 = \mu_1 m + \mu_2$ , and, thus,

$$m = \frac{1}{2}(\mu_1 + \sqrt{\mu_1^2 + 4\mu_2}). \tag{2.20}$$

The equation  $\hat{\mu}(z) = 1$  has one other root, viz.  $\gamma_1$  for which

$$\gamma_1^{-1} = -\frac{1}{2}(\sqrt{\mu_1^2 + 4\mu_2} - \mu_1). \tag{2.21}$$

Condition (B) is thus equivalent to  $|\gamma_1| > m^{-1/2}$ , or  $\gamma_1^{-2} < m$ ; by elementary algebra, this is equivalent to

$$\mu_1^3 + 3\mu_1\mu_2 + \mu_2 - \mu_2^2 > 0. \tag{2.22}$$

Thus, Theorem 2.1 applies when (2.22) holds, Theorem 2.2 when there is equality in (2.22), and Theorem 2.3 when the left-hand side of (2.22) is negative. (In this example, (2.16) is trivial.)

For a simple numerical example with  $\gamma_* = m^{-1/2}$ , take  $\mu_1 = 2$  and  $\mu_2 = 8$ . Then (2.20)–(2.21) yield  $m = 4$  and  $\gamma_1 = -\frac{1}{2}$ . We obtain, by (2.18), for example,

$$\frac{X_{n,1}}{\sqrt{nZ_n}} \xrightarrow{D} \zeta_1 \sim N\left(0, \frac{1}{768} \text{var}(N_2 - 2N_1)\right). \tag{2.23}$$

Suppose now instead that (2.22) holds, so Theorem 2.1 applies. Let  $\lambda := \gamma_1^{-1}$  be given by (2.21). Then  $1 - \hat{\mu}(z) = (1 - mz)(1 - \lambda z)$ , and, thus, (2.15) yields, for example,

$$\begin{aligned} \frac{m \text{var}\{\zeta_1\}}{m - 1} &= \oint_{|z|=m^{-1/2}} \frac{|z - m^{-1}|^2}{|1 - z|^2 |1 - \hat{\mu}(z)|^2} \Sigma(z) \frac{|dz|}{2\pi m^{-1/2}} \\ &= \frac{1}{m^2} \oint_{|z|=m^{-1/2}} \frac{\sigma_{11}|z|^2 + \sigma_{12}(z + \bar{z})|z|^2 + \sigma_{22}|z|^4}{|1 - z|^2 |1 - \lambda z|^2} \frac{|dz|}{2\pi m^{-1/2}}. \end{aligned}$$

This integral can be evaluated by expanding  $(1 - z)^{-1}(1 - \lambda z)^{-1}$  in a Taylor series, yielding

$$\text{var}(\zeta_1) = \frac{(m + \lambda)(\sigma_{11} + \sigma_{22}/m) + 2(1 + \lambda)\sigma_{12}}{m^2(m - \lambda)(m - \lambda^2)}.$$

**Remark 2.2.** By Theorem 2.1, the limit in (2.14) is degenerate only when the entire process is, and, thus, each  $X_{n,k}$  is degenerate. In contrast, the limit in (2.17) or the approximation in (2.19) can be degenerate even in other (special) situations. For example, let  $N_1$  be nondegenerate with  $\mathbb{E} N_1 = 2$ , let  $N_2 := 2N_1 + 4$ , and let  $N_k := 0$  for  $k > 2$ . Then  $\mu_1 = 2$  and  $\mu_2 = 8$ , and Example 2.2 shows that  $\gamma_* = \frac{1}{2} = m^{-1/2}$ ; further, (2.23) applies and yields  $X_{n,k}/\sqrt{nZ_n} \xrightarrow{D} 0$ .

We conjecture that in this case (and similar cases with  $\zeta_k = 0$  in Theorem 2.2),  $X_{n,k}/\sqrt{Z_n}$  has a nontrivial normal limit in distribution; we leave this as an open problem. Similarly, we conjecture that, when each  $\hat{\Xi}(\gamma_i)$  is degenerate in Theorem 2.3, the distribution of  $X_{n,k}$  is asymptotically determined by the next smallest roots in  $\Gamma_*$ .

**2.1. More notation**

For a random variable  $X$  in a Banach space  $\mathcal{B}$ , define  $\|X\|_{L^2(\mathcal{B})} := (\mathbb{E} \|X\|_{\mathcal{B}}^2)^{1/2}$ ; when  $\mathcal{B} = \mathbb{R}$  or  $\mathbb{C}$  abbreviate this to  $\|X\|_2$ . For infinite vectors  $\vec{x} = (x_j)_{j=0}^\infty$  and  $\vec{y} = (y_j)_{j=0}^\infty$ , let  $\langle \vec{x}, \vec{y} \rangle := \sum_{j=0}^\infty x_j y_j$ , assuming that the sum converges absolutely. Let  $C$  denote different constants that may depend on the distribution of the branching process (i.e. on the distribution of  $N$  and  $(\xi_i)$ ), but not on  $n$  and similar parameters; the constant may change from one occurrence to the next. We write  $O_{\text{a.s.}}(1)$  to denote a quantity that is bounded by a random constant that does not depend on  $n$ . All unspecified limits are as  $n \rightarrow \infty$ .

**3. Preliminaries**

Let

$$B_n := Z_n - Z_{n-1} \tag{3.1}$$

be the number of individuals born at time  $n$  (with  $B_0 = Z_0$ ). Thus,

$$Z_n = Z_{n-1} + B_n = \sum_{i=0}^n B_i, \quad n \geq 0. \tag{3.2}$$

Let  $B_{n,k}$  be the number of individuals born at time  $n + k$  to parents that are themselves born at time  $n$ , and, thus, are of age  $k$ . Then, recalling (A2),

$$B_n = \sum_{k=1}^n B_{n-k,k}, \quad n \geq 1. \tag{3.3}$$

Let  $\mathcal{F}_n$  be the  $\sigma$ -field generated by the life histories of all individuals born up to time  $n$ , with  $\mathcal{F}_n$  trivial for  $n < 0$ . Then  $B_{n,k}$  is  $\mathcal{F}_n$ -measurable, and  $B_n$  is  $\mathcal{F}_{n-1}$ -measurable by (3.3). Further,

$$\mathbb{E}(B_{n,k} \mid \mathcal{F}_{n-1}) = \mu_k B_n, \quad n \geq 0. \tag{3.4}$$

For  $k \geq 1$ , let

$$W_{n,k} := B_{n,k} - \mathbb{E}(B_{n,k} \mid \mathcal{F}_{n-1}) = B_{n,k} - \mu_k B_n, \tag{3.5}$$

so  $W_{n,k} = 0$  if  $n < 0$ . Then  $W_{n,k}$  is  $\mathcal{F}_n$ -measurable with

$$\mathbb{E}(W_{n,k} \mid \mathcal{F}_{n-1}) = 0. \tag{3.6}$$

Let further

$$W_n := B_n - \sum_{k=1}^n \mu_k B_{n-k} = B_n - \sum_{k=1}^{\infty} \mu_k B_{n-k}, \quad n \geq 0. \tag{3.7}$$

Thus,  $W_0 = B_0 = Z_0$ , and, for  $n \geq 1$ , by (3.3), (3.5), and (3.7),

$$W_n = \sum_{k=1}^n W_{n-k,k}. \tag{3.8}$$

**Lemma 3.1.** *Assume that (A1)–(A6) hold. Then, for all  $n \geq 1$  and  $k \geq 1$ ,  $\mathbb{E}[W_{n,k}^2] \leq Cr^{-2k}m^n$  and  $\mathbb{E}[W_n^2] \leq Cm^n$ .*

*Proof.* Recall that  $N_k$  is the number of children born to an individual of age  $k$ , and that  $\mathbb{E}N_k = \mu_k$ . From (2.2),  $\hat{\Xi}(r) \geq N_k r^k$  and, thus,

$$\text{var } N_k \leq \mathbb{E}N_k^2 \leq r^{-2k} \mathbb{E}[\hat{\Xi}(r)^2] = Cr^{-2k}. \tag{3.9}$$

Let  $n \geq 0$  and  $k \geq 1$ . Given  $\mathcal{F}_{n-1}$ ,  $B_{n,k}$  is the sum of  $B_n$  independent copies of  $N_k$ , and, thus (see (3.4), (3.5), and (3.9)),  $\mathbb{E}(W_{n,k}^2 \mid \mathcal{F}_{n-1}) = B_n \text{var}(N_k) \leq Cr^{-2k}B_n$ . Taking expectations and using (2.7),

$$\mathbb{E}[W_{n,k}^2] \leq Cr^{-2k} \mathbb{E}B_n \leq Cr^{-2k} \mathbb{E}Z_n \leq Cr^{-2k}m^n,$$

as asserted. Consequently,  $\|W_{n,k}\|_2 \leq Cr^{-k}m^{n/2}$  and, by (3.8) and Minkowski's inequality, using  $rm^{1/2} > 1$ ,

$$\|W_n\|_2 \leq \sum_{k=1}^n \|W_{n-k,k}\|_2 \leq Cm^{n/2} \sum_{k=1}^{\infty} (rm^{1/2})^{-k} \leq Cm^{n/2}. \quad \square$$

For  $n \geq 0$  and  $k \geq 1$ , by (2.10),

$$\begin{aligned} X_{n+1,k} &= Z_{n+1-k} - m^{-k}Z_{n+1} \\ &= X_{n,k-1} + m^{1-k}Z_n - m^{-k}Z_{n+1} \\ &= X_{n,k-1} + m^{-k}(mZ_n - Z_{n+1}). \end{aligned} \tag{3.10}$$



Also, for  $k \geq 0$ , by (3.1) and (2.10),

$$B_{n-k} = Z_{n-k} - Z_{n-k-1} = X_{n,k} - X_{n,k+1} + (m - 1)m^{-k-1}Z_n. \tag{3.11}$$

By (3.2), (3.7), and (3.11), recalling that  $X_{n,0} = 0$  by (2.10) and  $\hat{\mu}(m^{-1}) = 1$  by (A5), for  $n \geq 0$ ,

$$\begin{aligned} mZ_n - Z_{n+1} &= (m - 1)Z_n - B_{n+1} \\ &= (m - 1)Z_n - \sum_{k=1}^{\infty} \mu_k B_{n+1-k} - W_{n+1} \\ &= (m - 1)Z_n - \sum_{k=1}^{\infty} \mu_k (X_{n,k-1} - X_{n,k} + (m - 1)m^{-k}Z_n) - W_{n+1} \\ &= (m - 1)Z_n - \sum_{k=1}^{\infty} \mu_k (X_{n,k-1} - X_{n,k}) - (m - 1)\hat{\mu}(m^{-1})Z_n - W_{n+1} \\ &= \sum_{k=1}^{\infty} \mu_k (X_{n,k} - X_{n,k-1}) - W_{n+1}. \end{aligned}$$

Consequently, (3.10) yields, for  $n \geq 0$  and  $k \geq 1$ ,

$$X_{n+1,k} = X_{n,k-1} + m^{-k} \left( \sum_{j=1}^{\infty} \mu_j (X_{n,j} - X_{n,j-1}) - W_{n+1} \right). \tag{3.12}$$

Introduce the vector notation  $\vec{X}_n := (X_{n,k})_{k=0}^{\infty}$  and

$$\vec{v} := (0, m^{-1}, m^{-2}, \dots) = (m^{-k} \mathbf{1}\{k > 0\})_{k=0}^{\infty}, \tag{3.13}$$

and, for vectors  $\vec{y} = (y_k)_0^{\infty}$  such that the sum converges, define

$$\Psi((y_k)_0^{\infty}) := \sum_{k=1}^{\infty} \mu_k (y_k - y_{k-1}). \tag{3.14}$$

Let  $S$  be the shift operator  $S((y_k)_0^{\infty}) := (y_{k-1})_0^{\infty}$  with  $y_{-1} := 0$ , and let  $T$  be the linear operator

$$T(\vec{y}) := S(\vec{y}) + \Psi(\vec{y})\vec{v}. \tag{3.15}$$

Then (3.12) can be written, again recalling that  $X_{n,0} = 0$ ,

$$\vec{X}_{n+1} = S(\vec{X}_n) + (\Psi(\vec{X}_n) - W_{n+1})\vec{v} = T(\vec{X}_n) - W_{n+1}\vec{v}. \tag{3.16}$$

This recursion leads to the following formula.

**Lemma 3.2.** *For every  $n \geq 0$ ,*

$$\vec{X}_n = - \sum_{k=0}^n W_{n-k} T^k(\vec{v}). \tag{3.17}$$

*Proof.* For the initial value  $\vec{X}_0$ , (2.10) gives  $X_{0,k} = -m^{-k}Z_0$  for  $k \geq 1$ , and, thus, by (3.13),  $\vec{X}_0 = -Z_0\vec{v} = -W_0\vec{v}$ , recalling that  $W_0 = B_0 = Z_0$ . This verifies (3.17) for  $n = 0$ . The general case follows by (3.16) and induction.  $\square$

**Remark 3.1.** The proofs below show that the sum in (3.17) is dominated by the first few terms when  $\gamma_* > m^{-1/2}$ , and by the last few terms when  $\gamma_* < m^{-1/2}$ , while all terms are of about the same size when  $\gamma_* = m^{-1/2}$ . This explains much of the differences in behaviour seen in Section 2.

We now consider  $T$  defined in (3.15) as an operator on the complex Hilbert space  $\ell^2_R$  defined in (2.13) for suitable  $R > 0$ . Recall that the *spectrum*  $\sigma(T)$  of a linear operator in a complex Hilbert (or Banach) space is the set of complex numbers  $\lambda$  such that  $\lambda - T$  is *not* invertible (see, e.g. [4, Section VII.3]).

**Lemma 3.3.** *Suppose that  $1 \leq R < m$  and  $\hat{\mu}(R^{-1}) < \infty$ . Then  $\vec{v} \in \ell^2_R$ ,  $\Psi$  is a bounded linear functional on  $\ell^2_R$  and  $T$  is a bounded linear operator on  $\ell^2_R$ . Furthermore, if  $\lambda \in \mathbb{C}$  with  $|\lambda| > R$ , then  $\lambda \in \sigma(T)$  if and only if  $\lambda^{-1} \in \Gamma_*$ , i.e. if and only if  $\lambda \neq m$  and  $\hat{\mu}(\lambda^{-1}) = 1$ .*

*Proof.* By (2.13) and (3.13), and because  $R < m$ ,

$$\|\vec{v}\|_{\ell^2_R}^2 = \sum_{k=1}^{\infty} R^{2k} m^{-2k} < \infty.$$

Next, it is clear from (2.13) that the shift operator  $S$  is bounded on  $\ell^2_R$  (with norm  $R$ ). Further, by (2.1) and assumption,

$$\sum_{k=1}^{\infty} R^{-2k} \mu_k^2 \leq \hat{\mu}(R^{-1})^2 < \infty,$$

and it follows by the Cauchy–Schwarz inequality that  $\Psi_1((a_k)_0^\infty) := \sum_{k=1}^{\infty} \mu_k a_k$  defines a bounded linear functional  $\Psi_1$  on  $\ell^2_R$ . Since  $\Psi$  can be written  $\Psi = \Psi_1 - \Psi_1 S$ ,  $\Psi$  too is bounded. It now follows from (3.15) that  $T$  is a bounded linear operator on  $\ell^2_R$ .

For the final statement, note that the mapping  $(a_k)_0^\infty \mapsto \sum_{k=0}^{\infty} a_k z^k$  is an isometry of  $\ell^2_R$  onto the Hardy space  $H^2_R$  consisting of all functions  $f(z)$  analytic in the disc  $D_R$  such that

$$\|f\|_{H^2_R}^2 := \sup_{r < R} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta < \infty \tag{3.18}$$

(see, e.g. [5]). In particular,  $\vec{v}$  corresponds to the function

$$v(z) := \sum_{k=1}^{\infty} m^{-k} z^k = \frac{z/m}{1 - z/m} = \frac{z}{m - z}. \tag{3.19}$$

We use the same notation  $\Psi$ ,  $S$ , and  $T$  for the corresponding linear functional and operators on  $H^2_R$ , and note that the shift operator  $S$  on  $\ell^2_R$  corresponds to the multiplication operator  $Sf(z) = zf(z)$  on  $H^2_R$ . Definition (3.15) thus translates to

$$Tf(z) = zf(z) + \Psi(f)v(z). \tag{3.20}$$

Consequently, for any  $h \in H^2_R$ , the equation  $(\lambda - T)f = h$  is equivalent to

$$(\lambda - z)f(z) - \Psi(f)v(z) = h(z). \tag{3.21}$$

Any solution of (3.21) must be of the form

$$f(z) = c \frac{v(z)}{\lambda - z} + \frac{h(z)}{\lambda - z}, \tag{3.22}$$

where

$$c = \Psi(f) = c\Psi\left(\frac{v(z)}{\lambda - z}\right) + \Psi\left(\frac{h(z)}{\lambda - z}\right). \tag{3.23}$$

Suppose that  $|\lambda| > R$ ; then  $1/(\lambda - z)$  is a bounded analytic function on the domain  $D_R$ , so it follows from (3.18) and  $v, h \in H_R^2$  that  $v(z)/(\lambda - z) \in H_R^2$  and  $h(z)/(\lambda - z) \in H_R^2$ . If  $\Psi(v(z)/(\lambda - z)) \neq 1$  then (3.23) has a unique solution  $c$  for any  $h \in H_R^2$ , and, thus, (3.21) has a unique solution  $f \in H_R^2$ , given by (3.22), or in other words,  $\lambda - T$  is invertible on  $H_R^2$  and  $\lambda \notin \sigma(T)$ . (Continuity of  $(\lambda - T)^{-1}$  is automatic, by the closed graph theorem.)

Conversely, if  $\Psi(v(z)/(\lambda - z)) = 1$ , then (3.21) has either no solution or infinitely many solutions  $f$  for any given  $h \in H_R^2$ , and, therefore,  $\lambda \in \sigma(T)$ . Thus, for  $|\lambda| > R$ ,

$$\lambda \in \sigma(T) \iff \Psi\left(\frac{v(z)}{\lambda - z}\right) = 1. \tag{3.24}$$

We analyse the condition in (3.24) further. If  $|\lambda| > R$  and  $\lambda \neq m$ , then, by (3.19),

$$\frac{v(z)}{\lambda - z} = \frac{z}{(\lambda - z)(m - z)} = \frac{1}{m - \lambda} \left( \frac{\lambda}{\lambda - z} - \frac{m}{m - z} \right). \tag{3.25}$$

Now  $\lambda/(\lambda - z) = \sum_{k=0}^\infty \lambda^{-k} z^k$ , so, by (2.1) and (3.14),

$$\Psi\left(\frac{\lambda}{\lambda - z}\right) = \sum_{k=1}^\infty \mu_k \lambda^{-k} (1 - \lambda) = (1 - \lambda) \hat{\mu}(\lambda^{-1}).$$

Then, recalling  $\hat{\mu}(m^{-1}) = 1$  by (A5), (3.25) yields

$$\begin{aligned} \Psi\left(\frac{v(z)}{\lambda - z}\right) &= \frac{1}{m - \lambda} \left( \Psi\left(\frac{\lambda}{\lambda - z}\right) - \Psi\left(\frac{m}{m - z}\right) \right) \\ &= \frac{1}{m - \lambda} ((1 - \lambda) \hat{\mu}(\lambda^{-1}) - (1 - m) \hat{\mu}(m^{-1})) \\ &= \frac{1}{m - \lambda} ((1 - \lambda) \hat{\mu}(\lambda^{-1}) + m - 1). \end{aligned} \tag{3.26}$$

Consequently, for  $|\lambda| > R$  with  $\lambda \neq m$ , by (3.24) and (3.26),

$$\begin{aligned} \lambda \in \sigma(T) &\iff \Psi\left(\frac{v(z)}{\lambda - z}\right) = 1 \\ &\iff (1 - \lambda) \hat{\mu}(\lambda^{-1}) + m - 1 = m - \lambda \\ &\iff (1 - \lambda) \hat{\mu}(\lambda^{-1}) = 1 - \lambda \\ &\iff \hat{\mu}(\lambda^{-1}) = 1. \end{aligned}$$

For the special case  $\lambda = m$ , we let  $\lambda \rightarrow m$  in (3.26) and find by continuity that

$$\begin{aligned} \Psi\left(\frac{v(z)}{m - z}\right) &= \lim_{\lambda \rightarrow m} \Psi\left(\frac{v(z)}{\lambda - z}\right) = -\frac{d}{d\lambda} ((1 - \lambda) \hat{\mu}(\lambda^{-1})) \Big|_{\lambda=m} \\ &= \hat{\mu}(m^{-1}) - (m - 1) m^{-2} \hat{\mu}'(m^{-1}) \\ &< \hat{\mu}(m^{-1}) \\ &= 1, \end{aligned}$$

since  $\hat{\mu}'(x) > 0$  for  $x > 0$ . Hence,  $m \notin \sigma(T)$ . □

**Remark 3.2.** It is easily seen that  $\lambda \in \sigma(T)$  for every  $\lambda$  with  $|\lambda| \leq R$ —for example, take  $h = v$  in (3.21)–(3.22) and note that  $v(z)/(\lambda - z) \notin H^2_R$ . Thus, we have a complete description of the spectrum  $\sigma(T)$  on  $\ell^2_R$ .

**Lemma 3.4.** Suppose that  $1 \leq R < m$ ,  $\hat{\mu}(R^{-1}) < \infty$ , and  $\hat{\mu}(z) \neq 1$  for every complex  $z \neq m^{-1}$  with  $|z| < R^{-1}$ . Then, for every  $R_1 > R$ , there exists  $C = C(R_1)$  such that

$$\|T^n\|_{\ell^2_R} \leq CR_1^n, \quad n \geq 0. \tag{3.27}$$

*Proof.* By Lemma 3.3,  $T$  is a bounded linear operator on  $\ell^2_R$ . If  $\lambda \in \sigma(T)$  with  $|\lambda| > R$ , then  $\hat{\mu}(\lambda^{-1}) = 1$  and  $\lambda^{-1} \neq m^{-1}$ . By assumption, there is no such  $\lambda$ , so  $\sigma(T) \subseteq \{\lambda : |\lambda| \leq R\}$ . (In fact, equality holds by Remark 3.2.) In other words, the spectral radius

$$r(T) := \sup_{\lambda \in \sigma(T)} |\lambda| \leq R. \tag{3.28}$$

By the spectral radius formula [4, Lemma VII.3.4],  $r(T) = \lim_{n \rightarrow \infty} \|T^n\|^{1/n}$  and, thus, (3.28) implies that, for any  $R_1 > R$ ,  $\|T^n\|^{1/n} < R_1$  for large  $n$ , yielding (3.27).  $\square$

We use Lemma 3.4 when  $\gamma_* > m^{-1/2}$ , but when  $\gamma_* \leq m^{-1/2}$ , we use instead the following lemma, based on a more careful spectral analysis of  $T$ . Recall definitions (2.4)–(2.6).

**Lemma 3.5.** Let  $R = r^{-1} \geq 1$ , where  $\hat{\mu}(r) < \infty$ . Suppose also that  $\Gamma_{**} = \{\gamma_1, \dots, \gamma_q\} \neq \emptyset$ , and that (2.16) holds. Let  $\lambda_i := \gamma_i^{-1}$ . For  $i = 1, \dots, q$ , there exist eigenvectors  $\vec{v}_i$  with  $T\vec{v}_i = \lambda_i\vec{v}_i$ , linear projections  $P_i$  with range  $\mathcal{R}(P_i) = \{c\vec{v}_i : c \in \mathbb{C}\}$  (i.e. the span of  $\vec{v}_i$ ), a bounded operator  $T_0$  in  $\ell^2_R$ , and a constant  $R < \gamma_*^{-1}$  such that, for any  $n \geq 0$ ,

$$T^n = T_0^n + \sum_{i=1}^q \lambda_i^n P_i \tag{3.29}$$

and

$$\|T_0^n\|_{\ell^2_R} \leq C\tilde{R}^n. \tag{3.30}$$

Explicitly,

$$\vec{v}_i = P_i(\vec{v}) = \frac{1}{\gamma_i(\gamma_i - 1)\hat{\mu}'(\gamma_i)} (\gamma_i^k - m^{-k})_k. \tag{3.31}$$

*Proof.* Since the points in  $\Gamma_*$  are isolated, there exists  $\tilde{r} > \gamma_*$  such that  $|z| > \tilde{r}$  for any  $z \in \Gamma_* \setminus \Gamma_{**}$ . We may assume that  $\tilde{r} < r$ . Let  $\tilde{R} := \tilde{r}^{-1} > R$ . By Lemma 3.3,  $\lambda_i = \gamma_i^{-1} \in \sigma(T)$  with  $|\lambda_i| = \gamma_i^{-1}$ , and  $|\lambda| < \tilde{R} < \gamma_*^{-1}$  for any  $\lambda \in \sigma(T) \setminus \{\lambda_1, \dots, \lambda_q\}$ .

Then, since  $\lambda_1, \dots, \lambda_q$  are isolated points in  $\sigma(T)$ , by standard functional calculus (see, e.g. [4, Section VII.3]), there exist commuting projections (not necessarily orthogonal)  $P_0, \dots, P_q$  in  $\ell^2_R$  such that  $\sum_{i=0}^q P_i = 1$ ,  $T$  maps each subspace  $E_i := P_i(\ell^2_R)$  into itself, and if  $\hat{T}_i$  is the restriction of  $T$  to  $E_i$ , then  $\hat{T}_i$  has spectrum  $\sigma(\hat{T}_i) = \{\lambda_i\}$  for  $1 \leq i \leq q$  and  $\sigma(\hat{T}_0) = \sigma(T) \setminus \{\lambda_i\}_1^q$ . In particular, the spectral radius  $r(\hat{T}_0) < \tilde{R}$ , and, thus, by the spectral radius formula [4, Lemma VII.3.4],

$$\|\hat{T}_0^n\| \leq C\tilde{R}^n, \quad n \geq 0.$$

Let  $T_0 := TP_0$ . Then  $T_0^n = T^n P_0 = \hat{T}_0^n P_0$ , and (3.30) follows.

It remains to show that the spaces  $E_i = \mathcal{R}(P_i)$  are one dimensional, and spanned by the vectors  $\vec{v}_i$  in (3.31). As in the proof of Lemma 3.3, we use the isometry  $(a_k)_0^\infty \mapsto \sum_{k=0}^\infty a_k z^k$  of  $\ell^2_R$  onto  $H^2_R$ .

For each  $\lambda_i$ ,  $\hat{\mu}(\lambda_i^{-1}) = 1$ , so  $\Psi(v(z)/(\lambda_i - z)) = 1$  by (3.26) (see also (3.24)). Then taking  $h = 0$  in (3.21)–(3.23), conclude that the kernel  $\mathcal{N}(\lambda_i - T)$  is one dimensional and spanned by  $v(z)/(\lambda_i - z)$ . Similarly, again by (3.21)–(3.23), the range  $\mathcal{R}(\lambda_i - T)$  is given by

$$\mathcal{R}(\lambda_i - T) = \left\{ h \in \ell_R^2 : \Psi\left(\frac{h(z)}{\lambda_i - z}\right) = 0 \right\}. \tag{3.32}$$

Differentiating (3.26), we find, for  $|\lambda| > R$  with  $\lambda^{-1} \in \Gamma_*$ , i.e.  $\lambda \neq m$  and  $\hat{\mu}(\lambda^{-1}) = 1$ ,

$$\begin{aligned} \Psi\left(\frac{v(z)}{(\lambda - z)^2}\right) &= -\frac{d}{d\lambda} \Psi\left(\frac{v(z)}{\lambda - z}\right) \\ &= \frac{d}{d\lambda} \left(1 - \Psi\left(\frac{v(z)}{\lambda - z}\right)\right) \\ &= \frac{d}{d\lambda} \frac{(1 - \lambda)(1 - \hat{\mu}(\lambda^{-1}))}{m - \lambda} \\ &= \frac{(1 - \lambda)\hat{\mu}'(\lambda^{-1})}{(m - \lambda)\lambda^2}. \end{aligned} \tag{3.33}$$

Thus, (2.16) implies that  $\Psi(v(z)/(\lambda_i - z)^2) \neq 0$ , and so  $v(z)/(\lambda_i - z) \notin \mathcal{R}(\lambda_i - T)$  by (3.32). Hence,  $\mathcal{N}(\lambda_i - T) \cap \mathcal{R}(\lambda_i - T) = \{0\}$ . Consequently, for every  $h \in \mathcal{R}(\lambda_i - T)$ , (3.21) has a unique solution  $f \in \mathcal{R}(\lambda_i - T)$ , i.e. the restriction of  $\lambda_i - T$  to  $\mathcal{R}(\lambda_i - T)$  is invertible.

It follows that  $P_i$  is the projection onto  $\mathcal{N}(\lambda_i - T) = \{cv(z)/(\lambda_i - z)\}$  that vanishes on  $\mathcal{R}(\lambda_i - T)$ , so, by (3.32),  $P_i$  is given by

$$P_i(f(z)) = \frac{\Psi(f(z)/(\lambda_i - z))}{\Psi(v(z)/(\lambda_i - z)^2)} \frac{v(z)}{\lambda_i - z}. \tag{3.34}$$

In particular, since  $\Psi(v(z)/(\lambda_i - z)) = 1 \neq 0$ , we see that  $P_i(v)$  is a nonzero multiple of  $v(z)/(\lambda_i - z)$ . Let  $\vec{v}_i := P_i(\vec{v})$ . Thus,  $T\vec{v}_i = \lambda_i\vec{v}_i$ , and, for  $n \geq 0$ ,

$$T^n = T^n P_0 + \sum_{i=1}^q T^n P_i = T_0^n + \sum_{i=1}^q \lambda_i^n P_i,$$

showing (3.29). Finally, (3.33) and (3.34) yield

$$v_i(z) := P_i(v(z)) = \frac{(m - \lambda_i)\lambda_i^2}{(1 - \lambda_i)\hat{\mu}'(\lambda_i^{-1})} \frac{v(z)}{\lambda_i - z},$$

and (3.31) follows because  $\lambda_i = \gamma_i^{-1}$  and, by (3.19), for  $|\lambda| > R$ ,

$$(m - \lambda) \frac{v(z)}{\lambda - z} = \frac{\lambda}{\lambda - z} - \frac{m}{m - z} = \sum_{k=0}^{\infty} (\lambda^{-k} - m^{-k})z^k. \tag{3.35}$$

This completes the proof. □

**Remark 3.3.** It follows also that (2.16) implies that the points  $\lambda_i \in \sigma(T)$  are simple poles of the resolvent  $(\lambda - T)^{-1}$ , and conversely. Lemma 3.5 can be extended without assuming that (2.16) holds; the general result is similar but more complicated, and is left to the reader (cf. [4, Theorem VII.3.18]).

We also use another similar calculation.

**Lemma 3.6.** *Suppose that  $1 \leq R < m$  and  $\hat{\mu}(R^{-1}) < \infty$ . If  $|\lambda| > R$  and  $\hat{\mu}(\lambda^{-1}) \neq 1$ , then*

$$(\lambda - T)^{-1}(\vec{v}) = \frac{1}{(1 - \lambda)(1 - \hat{\mu}(\lambda^{-1}))}(\lambda^{-k} - m^{-k})_k.$$

*Proof.* Let  $h = v$  in (3.21)–(3.23); then

$$(\lambda - T)^{-1}v(z) = f(z) = b \frac{v(z)}{\lambda - z}$$

for a constant  $b$  such that  $b = \Psi(f) + 1$ . Using (3.26), this yields

$$b - 1 = \Psi(f) = \frac{b}{m - \lambda}((1 - \lambda)\hat{\mu}(\lambda^{-1}) + m - 1)$$

with the solution

$$b = \frac{m - \lambda}{(1 - \lambda)(1 - \hat{\mu}(\lambda^{-1}))}.$$

Then, for  $|z| < R$ , on using (3.35),

$$f(z) = b \frac{v(z)}{\lambda - z} = \frac{1}{(1 - \lambda)(1 - \hat{\mu}(\lambda^{-1}))} \sum_{k=0}^{\infty} (\lambda^{-k} - m^{-k})z^k. \quad \square$$

#### 4. A first normal convergence result

Let  $\vec{\eta} := (\eta_0, \eta_1, \eta_2, \dots)$ , where  $(\eta_k)_0^\infty$  are jointly normal random variables with means  $\mathbb{E} \eta_k = 0$  and covariances (cf. (2.11))

$$\text{cov}(\eta_j, \eta_k) = \sigma_{jk} = \text{cov}(N_j, N_k). \tag{4.1}$$

Note that, since  $N_0 = 0, \eta_0 = 0$ .

**Lemma 4.1.** *Assume that (A1)–(A6) hold, and let  $\vec{\eta}^{(k)} = (\eta_j^{(k)})_{j=0}^\infty, k = 1, 2, \dots$ , be independent copies of the random vector  $\eta$ . Then, as  $n \rightarrow \infty$ , jointly for all  $(j, k)$  with  $j \geq 0$  and  $k \geq 0$ ,*

$$Z_n^{-1/2} W_{n-k,j} \xrightarrow{D} \left(1 - \frac{1}{m}\right)^{1/2} m^{-k/2} \eta_j^{(k)}. \tag{4.2}$$

*Proof.* First fix  $k \geq 0$ . Given  $B_{n-k}$ , the vector  $\vec{B}_{n-k} := (B_{n-k,j})_{j=0}^\infty$  is the sum of  $B_{n-k}$  independent copies of the random vector  $\vec{N}$ , and, by (3.5), the vector  $\vec{W}_{n-k} := (W_{n-k,j})_{j=0}^\infty$  is the sum of  $B_{n-k}$  independent copies of the centered random vector  $\vec{N} - \mathbb{E} \vec{N}$ . By (3.1) and (2.9),

$$\frac{B_n}{Z_n} = 1 - \frac{Z_{n-1}}{Z_n} \xrightarrow{\text{a.s.}} 1 - m^{-1} > 0. \tag{4.3}$$

In particular,  $B_n \rightarrow \infty$  a.s., and, thus,  $B_{n-k} \rightarrow \infty$ . Then, by the central limit theorem for i.i.d. finite-dimensional vector-valued random variables and the definition of  $\eta_j$ ,

$$B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{D} \eta_j \stackrel{D}{=} \eta_j^{(k)}, \tag{4.4}$$

jointly for any finite set of  $j \geq 0$ . Moreover, by (2.9) and (4.3),

$$\frac{B_{n-k}}{Z_n} \xrightarrow{\text{a.s.}} \left(1 - \frac{1}{m}\right) m^{-k}, \tag{4.5}$$

so (4.2) for fixed  $k$  follows from (4.3) and (4.4).

To extend this result to several  $k$ , we must take account of the fact that  $W_{n-k,j}$  for different  $k$  are, in general, dependent. For example, conditioned on  $Z_{n-1}$  and  $B_{n-1}$ ,  $W_{n-1,1}$  determines  $B_{n-1,1}$  which contributes to  $B_n$ , and, thus, influences  $W_{n,j}$ . We therefore approximate  $W_{n-k,j}$  as follows.

We can assume that, for each  $k$ , we have an infinite sequence  $(\vec{N}^{(k,i)})_{i \geq 1}$  of independent copies of  $\vec{N}$ , such that  $\vec{W}_{n-k}$  is the sum  $\sum_{i=1}^{B_{n-k}} \vec{N}_j^{(k,i)}$  of the first  $B_{n-k}$  vectors; furthermore, these sequences for different  $k$  are independent.

Fix  $J, K \geq 1$  and consider only  $j \leq J$  and  $k \leq K$ . For  $0 \leq k \leq K$ , define

$$\begin{aligned} \bar{B}_{n-k} &:= \lfloor m^{K-k} B_{n-K} \rfloor, \\ \bar{W}_{n-k,j} &:= \sum_{i=1}^{\bar{B}_{n-k}} \vec{N}_j^{(k,i)}. \end{aligned} \tag{4.6}$$

Then, by the central limit theorem, exactly as for (4.4), jointly for all  $j \leq J$  and  $k \leq K$ ,

$$\bar{B}_{n-k}^{-1/2} \bar{W}_{n-k,j} \xrightarrow{D} \eta_j^{(k)}. \tag{4.7}$$

If here we condition on  $B_{n-K}$ , the left-hand sides for different  $k$  are independent; also, by (4.3) and (2.9),  $\bar{B}_{n-k}/B_{n-k} \xrightarrow{\text{a.s.}} 1$  for every  $k$ . Hence, (4.7) yields, jointly,

$$B_{n-k}^{-1/2} \bar{W}_{n-k,j} \xrightarrow{D} \eta_j^{(k)}. \tag{4.8}$$

Moreover, using (4.6),

$$\mathbb{E}((\bar{W}_{n-k,j} - W_{n-k,j})^2 \mid B_{n-k}, \bar{B}_{n-k}) = |B_{n-k} - \bar{B}_{n-k}| \text{var } N_j,$$

and, consequently, for every fixed  $j \geq 0, k \geq 0$  and  $\varepsilon > 0$ ,

$$\mathbb{P}(|\bar{W}_{n-k,j} - W_{n-k,j}| > \varepsilon B_{n-k}^{1/2} \mid B_{n-k}, \bar{B}_{n-k}) \leq \left| 1 - \frac{\bar{B}_{n-k}}{B_{n-k}} \right| \sigma_{jj} \varepsilon^{-2} \xrightarrow{\text{a.s.}} 0.$$

Taking expectations, we obtain by dominated convergence that, for every  $j$  and  $k$ ,  $\mathbb{P}(|\bar{W}_{n-k,j} - W_{n-k,j}| > \varepsilon B_{n-k}^{1/2}) \rightarrow 0$  for every  $\varepsilon > 0$ , and, thus,

$$B_{n-k}^{-1/2} \bar{W}_{n-k,j} - B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{P} 0. \tag{4.9}$$

Combining (4.8) and (4.9) yields, still jointly for all  $j \leq J$  and  $k \leq K$ ,

$$B_{n-k}^{-1/2} W_{n-k,j} \xrightarrow{D} \eta_j^{(k)}.$$

The result follows from this and (4.5), since  $J$  and  $K$  are arbitrary. □

### 5. First proof of Theorem 2.1

In this section we assume that (A1)–(A6) hold and also (B), i.e.  $\gamma_* > m^{-1/2}$ . In other words (cf. (2.5)), each  $z \in \Gamma_*$  satisfies  $|z| > m^{-1/2}$ . Hence, we can decrease  $r$  so that the disc  $D_r$  contains no roots of  $\hat{\mu}(z) = 1$  except  $m^{-1}$ , and still  $r > m^{-1/2}$ . Then, with  $R := 1/r$  and assuming that (A1)–(A6) hold,  $\gamma_* > m^{-1/2}$  is equivalent to the following condition.

- (B') There exists  $R$  satisfying  $1 \leq R < m^{1/2}$  such that  $\hat{\mu}(R^{-1}) < \infty$  and  $\hat{\mu}(z) \neq 1$  for every complex  $z \neq m^{-1}$  for which  $|z| < R^{-1}$ .

Fix  $R$  such that (B') holds and (A6) holds with  $r = 1/R$ ; note that  $R$  may be chosen arbitrarily close to  $m^{1/2}$ . Fix also  $R_1$  with  $R < R_1 < m^{1/2}$ . Then (B') and Lemma 3.4 show that (3.27) holds, i.e.  $\|T^n\|_{\ell_R^2} = O(R_1^n)$ .

**Lemma 5.1.** *Assume that (A1)–(A6) and (B) hold. If  $R < m^{1/2}$  then*

$$\mathbb{E} \|\vec{X}_n\|_{\ell_R^2}^2 \leq Cm^n, \tag{5.1}$$

so, for all  $n, k \geq 0$ ,

$$\mathbb{E} X_{n,k}^2 \leq CR^{-2k}m^n. \tag{5.2}$$

*Proof.* By (3.17), Lemma 3.1, (3.27), and Minkowski's inequality,

$$\|\vec{X}_n\|_{L^2(\ell_R^2)} \leq \sum_{k=0}^n \|W_{n-k}\|_{L^2} \|T^k(\vec{v})\|_{\ell_R^2} \leq C \sum_{k=0}^n m^{(n-k)/2} R_1^k = Cm^{n/2} \sum_{k=0}^{\infty} \left(\frac{R_1}{m^{1/2}}\right)^k = Cm^{n/2}.$$

This yields (5.1), and (5.2) follows by (2.13). □

For convenience, define  $W_{n,j}$  for  $n < 0$  by  $W_{-1,1} := W_0$  and  $W_{n,j} = 0$  for  $n \leq -1$  and  $j \geq 1$  with  $(n, j) \neq (-1, 1)$ . Then (3.8) holds also for  $n \leq 0$ , provided the sum is extended to  $\infty$ , and (3.17) can be written as

$$\vec{X}_n = - \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} W_{n-k-j,j} T^k(\vec{v}). \tag{5.3}$$

For each finite  $M$ , define also the truncated sum

$$\vec{X}_{n,M} := - \sum_{k=0}^M \sum_{j=1}^M W_{n-k-j,j} T^k(\vec{v}). \tag{5.4}$$

Lemma 4.1 implies that, for any fixed  $M$ , as  $n \rightarrow \infty$ ,

$$Z_n^{-1/2} \vec{X}_{n,M} \xrightarrow{D} - \sum_{k=0}^M \sum_{j=1}^M (1 - m^{-1})^{1/2} m^{-(k+j)/2} \eta_j^{(k+j)} T^k(\vec{v}) \quad \text{in } \ell_R^2. \tag{5.5}$$

Furthermore, by (5.3)–(5.4), Minkowski's inequality, Lemma 3.1, and (3.27), regarding  $\vec{X}_n$  and  $\vec{X}_{n,M}$  as elements of  $L^2(\ell_R^2)$ , the space of  $\ell_R^2$ -valued random variables with square integrable norm,

$$\begin{aligned} \|\vec{X}_n - \vec{X}_{n,M}\|_{L^2(\ell_R^2)} &\leq \sum_{k>M \text{ or } j>M} \|W_{n-k-j,j}\|_{L^2} \|T^k(\vec{v})\|_{\ell_R^2} \\ &\leq C \sum_{k>M \text{ or } j>M} r^{-j} m^{(n-k-j)/2} R_1^k \\ &= Cm^{n/2} \sum_{k>M \text{ or } j>M} \left(\frac{R}{m^{1/2}}\right)^j \left(\frac{R_1}{m^{1/2}}\right)^k. \end{aligned} \tag{5.6}$$

Since the sum on the right-hand side of (5.6) converges, it tends to 0 as  $M \rightarrow \infty$ , and, thus,  $m^{-n/2}(\vec{X}_n - \vec{X}_{n,M}) \rightarrow 0$  in  $L^2(\ell_R^2)$ , hence in probability, uniformly in  $n$ . Since  $Z_n/m^n \xrightarrow{\text{a.s.}} \mathcal{Z} > 0$  (see (2.8)),  $\sup_n m^n/Z_n$  is an a.s. finite random variable; hence, also

$$Z_n^{-1/2}(\vec{X}_n - \vec{X}_{n,M}) = \left(\frac{m^n}{Z_n}\right)^{1/2} m^{-n/2}(\vec{X}_n - \vec{X}_{n,M}) \xrightarrow{P} 0 \quad \text{as } M \rightarrow \infty, \text{ uniformly in } n.$$



Also, the right-hand side of (5.5) converges as  $M \rightarrow \infty$  in  $L^2(\ell_R^2)$ , hence in distribution, since, by (3.9)

$$\mathbb{E}[(\eta_j^{(k)})^2] = \text{var } N_j \leq Cr^{-2j} = CR^{2j},$$

and, thus, using also (3.27),

$$\begin{aligned} \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} m^{-(k+j)/2} \|\eta_j^{(k+j)} T^k(\vec{v})\|_{L^2(\ell_R^2)} &= \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} m^{-(k+j)/2} \|\eta_j^{(k+j)}\|_{L^2} \|T^k(\vec{v})\|_{\ell_R^2} \\ &\leq C \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} m^{-(k+j)/2} R^j R_1^k \\ &< \infty. \end{aligned} \tag{5.7}$$

It follows (see [2, Theorem 4.2]) that (5.5) extends to  $M = \infty$ , i.e. as  $n \rightarrow \infty$ ,

$$Z_n^{-1/2} \vec{X}_n \xrightarrow{D} - \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} (1 - m^{-1})^{1/2} m^{-(k+j)/2} \eta_j^{(k+j)} T^k(\vec{v}) \quad \text{in } \ell_R^2. \tag{5.8}$$

The right-hand side is obviously a Gaussian random vector in  $\ell_R^2$ , which we write as  $\vec{\zeta} = (\zeta_0, \zeta_1, \dots)$ . Then (5.8) yields (2.14).

It remains to calculate the covariances of  $\zeta_k$ . Let  $\vec{a} = (a_0, a_1, \dots)$  be a (real) vector with only finitely many nonzero elements. Then, by (5.8),

$$\sum_{\ell=0}^{\infty} a_{\ell} \zeta_{\ell} = \langle \vec{a}, \vec{\zeta} \rangle = -(1 - m^{-1})^{1/2} \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} m^{-(k+j)/2} \eta_j^{(k+j)} \langle T^k(\vec{v}), \vec{a} \rangle, \tag{5.9}$$

where the sum converges absolutely in  $L^2$  by (5.7). From the definition of  $\eta_j^{(k)}$  in (4.1) and Lemma 4.1,

$$\text{cov}\left(m^{-k/2} \eta_i^{(k)}, m^{-\ell/2} \eta_j^{(\ell)}\right) = m^{-(k+\ell)/2} \delta_{k,\ell} \sigma_{ij} = \oint_{|w|=m^{-1/2}} \sigma_{ij} w^k \bar{w}^{\ell} \frac{|dw|}{2\pi m^{-1/2}}.$$

Hence, (5.9) yields

$$\begin{aligned} \frac{\text{var}(\langle \vec{a}, \vec{\zeta} \rangle)}{1 - m^{-1}} &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \langle T^k(\vec{v}), \vec{a} \rangle \langle T^{\ell}(\vec{v}), \vec{a} \rangle \oint_{|w|=m^{-1/2}} \sigma_{ij} w^{k+i} \bar{w}^{\ell+j} \frac{|dw|}{2\pi m^{-1/2}} \\ &= \oint_{|w|=m^{-1/2}} \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \sigma_{ij} w^i \bar{w}^j \left| \sum_{k=0}^{\infty} w^k \langle T^k(\vec{v}), \vec{a} \rangle \right|^2 \frac{|dw|}{2\pi m^{-1/2}}. \end{aligned} \tag{5.10}$$

Furthermore, if  $|w| = m^{-1/2}$  then  $\sum_{k=0}^{\infty} \|w^k T^k(\vec{v})\|_{\ell_R^2} < \infty$  by (3.27), and, thus,

$$\sum_{k=0}^{\infty} w^k T^k(\vec{v}) = (1 - wT)^{-1}(\vec{v}). \tag{5.11}$$

Set  $\lambda := w^{-1}$ , so  $|\lambda| = m^{1/2} > R$ . As in the proof of Lemma 3.3, we use the standard isometry  $\ell_R^2 \rightarrow H_R^2$ , and let  $f(z) \in H_R^2$  be the function corresponding to  $(1 - wT)^{-1}(\vec{v}) = \lambda(\lambda - T)^{-1}(\vec{v})$ . Then (cf. (3.20)–(3.21))

$$(\lambda - z)f(z) - \Psi(f)v(z) = (\lambda - T)f(z) = \lambda v(z);$$

hence (cf (3.21)–(3.23)),

$$f(z) = b \frac{v(z)}{\lambda - z}$$

for a constant  $b$  such that  $b = \Psi(f) + \lambda$ . By (3.26), this yields

$$b - \lambda = \Psi(f) = \frac{b}{m - \lambda} ((1 - \lambda)\hat{\mu}(\lambda^{-1}) + m - 1),$$

which has the solution

$$b = \frac{\lambda(m - \lambda)}{(1 - \lambda)(1 - \hat{\mu}(\lambda^{-1}))}.$$

Then, using (3.25), for  $|z| \leq R$ ,

$$\begin{aligned} f(z) &= b \frac{v(z)}{\lambda - z} \\ &= \frac{\lambda}{(1 - \lambda)(1 - \hat{\mu}(\lambda^{-1}))} \left( \frac{\lambda}{\lambda - z} - \frac{m}{m - z} \right) \\ &= \frac{\lambda}{(1 - \lambda)(1 - \hat{\mu}(\lambda^{-1}))} \sum_{\ell=0}^{\infty} (\lambda^{-\ell} - m^{-\ell}) z^{\ell}. \\ &= \frac{1}{(w - 1)(1 - \hat{\mu}(w))} \sum_{\ell=0}^{\infty} (w^{\ell} - m^{-\ell}) z^{\ell}. \end{aligned}$$

Thus,  $(1 - wT)^{-1}(\vec{v}) = (((w - 1)(1 - \hat{\mu}(w)))^{-1}(w^{\ell} - m^{-\ell}))_{\ell}$  and, using (5.11),

$$\sum_{k=0}^{\infty} w^k \langle T^k(\vec{v}), \vec{a} \rangle = \langle (1 - wT)^{-1}(\vec{v}), \vec{a} \rangle = \frac{1}{(w - 1)(1 - \hat{\mu}(w))} \sum_{\ell=0}^{\infty} a_{\ell}(w^{\ell} - m^{-\ell}). \tag{5.12}$$

Hence, (2.15) follows from (5.10).

Finally, by (2.15), the variable  $\zeta_k$  is degenerate only if  $\Sigma(z) = 0$  for every  $z$  with  $|z| = m^{-1/2}$ , and, thus, by (2.12),  $\hat{\Xi}(z) = \hat{\mu}(z)$  a.s. for every such  $z$ , which, by (2.1)–(2.2), implies that  $N_k = \mu_k$  a.s. for every  $k$ .

### 6. A martingale

In the remaining sections, we let  $R := r^{-1} < m^{1/2}$ , where  $r$  is as in (A6), and, by decreasing  $r$ ,  $R$  can be arbitrarily close to  $m^{1/2}$ . Consider again the operator  $T$  on  $\ell^2_R$ .

Fix a real vector  $\vec{a} \in \ell^2_{R^{-1}}$  (for example, any finite real vector), and write

$$\alpha_k = \alpha_k(\vec{a}) := \langle T^k(\vec{v}), \vec{a} \rangle. \tag{6.1}$$

Then (3.17) and (3.8) yield

$$\langle \vec{X}_n, \vec{a} \rangle = - \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} W_{n-k-j, j} \alpha_k = - \sum_{\ell=0}^n \sum_{j=1}^{n-\ell} W_{\ell, j} \alpha_{n-j-\ell}. \tag{6.2}$$

Define

$$\Delta M_{n, \ell} := \sum_{j=1}^{n-\ell} \alpha_{n-j-\ell} W_{\ell, j}, \tag{6.3}$$

$$M_{n, k} := \sum_{\ell=0}^k \Delta M_{n, \ell}. \tag{6.4}$$

Then (3.6) shows that  $\mathbb{E}(\Delta M_{n,\ell} \mid \mathcal{F}_{\ell-1}) = 0$ , and, thus,  $(M_{n,k})_{k=0}^n$  is a martingale with respect to  $(\mathcal{F}_k)_k$ . Also, by (6.2),

$$\langle \vec{X}_n, \vec{a} \rangle = -M_{n,n}. \tag{6.5}$$

Conditioned on  $\mathcal{F}_{\ell-1}$ , the vector  $(W_{\ell,j})_j$  is the sum of  $B_\ell$  independent copies of  $\vec{N} - \mathbb{E}\vec{N}$ , where  $\vec{N} = (N_j)_0^\infty$ , and, thus, recalling (2.11),

$$Q_{n,\ell} := \mathbb{E}((\Delta M_{n,\ell})^2 \mid \mathcal{F}_{\ell-1}) = B_\ell \operatorname{var} \left( \sum_{j=1}^{n-\ell} \alpha_{n-\ell-j} N_j \right) = B_\ell \sum_{i,j=1}^{n-\ell} \sigma_{ij} \alpha_{n-\ell-i} \alpha_{n-\ell-j}. \tag{6.6}$$

The conditional quadratic variation of the martingale  $(M_{n,k})_k$  is thus

$$V_n := \sum_{\ell=0}^n Q_{n,\ell} = \sum_{\ell=0}^n B_\ell \sum_{i,j=1}^{n-\ell} \sigma_{ij} \alpha_{n-\ell-i} \alpha_{n-\ell-j} = \sum_{\ell=0}^n B_{n-\ell} \sum_{i,j=1}^{\ell} \sigma_{ij} \alpha_{\ell-i} \alpha_{\ell-j}. \tag{6.7}$$

By (2.2),  $N_k \leq r^{-k} \hat{\mathbb{E}}(r)$ , and, thus, by (2.11) and the Cauchy–Schwarz inequality,

$$|\sigma_{ij}| \leq r^{-i-j} \mathbb{E}[\hat{\mathbb{E}}(r)^2] = CR^{i+j}. \tag{6.8}$$

### 7. Second proof of Theorem 2.1

We give here another proof of Theorem 2.1, based on a martingale central limit theorem and the martingale in Section 6, because with small modifications this new proof also applies to Theorem 2.2 (see Section 8) and we prefer to present it first for Theorem 2.1. (The proof in Section 5 does not seem to extend easily to Theorem 2.2.)

Let  $R$  and  $R_1$  be as in Section 5. Then (3.27) and (6.1) show that, for fixed  $\vec{a}$  and with  $C = C(\vec{a})$ ,

$$|\alpha_k| \leq CR_1^k. \tag{7.1}$$

Consequently, by (6.6), (6.8), and (7.1), since  $R/R_1 < 1$ ,

$$\frac{Q_{n,\ell}}{B_\ell} = \sum_{i,j=1}^{n-\ell} \sigma_{ij} \alpha_{n-\ell-i} \alpha_{n-\ell-j} \leq C \sum_{i,j=1}^{\infty} R^{i+j} R_1^{2(n-\ell)-i-j} \leq CR_1^{2(n-\ell)}. \tag{7.2}$$

Hence, by (6.7), (6.6), (3.1) and (2.9), using dominated convergence justified by (7.2) and  $R_1^2/m < 1$ ,

$$\begin{aligned} \frac{V_n}{Z_n} &= \sum_{\ell=0}^n \frac{B_{n-\ell}}{Z_n} \frac{Q_{n,n-\ell}}{B_{n-\ell}} \\ &= \sum_{\ell=0}^n \frac{Z_{n-\ell} - Z_{n-\ell-1}}{Z_n} \sum_{i,j=1}^{\ell} \sigma_{ij} \alpha_{\ell-i} \alpha_{\ell-j} \\ &\xrightarrow{\text{a.s.}} \sigma^2(\vec{a}) \\ &:= \sum_{\ell=0}^{\infty} (m^{-\ell} - m^{-\ell-1}) \sum_{i,j=1}^{\ell} \sigma_{ij} \alpha_{\ell-i} \alpha_{\ell-j}. \end{aligned} \tag{7.3}$$

We cannot use a martingale central limit theorem directly for the martingale  $(M_{n,k})_k$  defined in (6.4), because the calculations above show that most of the conditional quadratic variation  $V_n$  comes from the last few terms (cf. Remark 3.1). We therefore introduce another martingale.

Number the individuals 1, 2, . . . in order of birth, with arbitrary order at ties, and let  $\mathcal{G}_\ell$  be the  $\sigma$ -field generated by the life histories of individuals 1, . . . ,  $\ell$ . Each  $Z_n$  is a stopping time with respect to  $(\mathcal{G}_\ell)_\ell$ , and  $\mathcal{G}_{Z_n} = \mathcal{F}_n$ .

We refine the martingale  $(M_{n,k})_k$  by adding the contribution from each individual separately. Let  $\tau_i$  denote the birth time of  $i$ , and  $N_{i,k}$  the copy of  $N_k$  for  $i$  (i.e. the number of children born to  $i$  at age  $k$ ). Let

$$\begin{aligned} \Delta \hat{M}_{n,i} &:= \sum_{j=1}^{n-\tau_i} \alpha_{n-\tau_i-j} (N_{i,j} - \mu_j), \\ \hat{M}_{n,k} &:= \sum_{i=1}^k \Delta \hat{M}_{n,i}. \end{aligned} \tag{7.4}$$

Then  $(\hat{M}_{n,k})_k$  is a  $(\mathcal{G}_k)_k$ -martingale with  $\hat{M}_{n,\infty} = \hat{M}_{n,Z_n} = M_{n,n} = -\langle \bar{X}_n, \bar{a} \rangle$  (see (6.3)–(6.5)), and the conditional quadratic variation

$$\hat{V}_n := \sum_i \mathbb{E}((\Delta \hat{M}_{n,i})^2 \mid \mathcal{G}_{i-1}) = V_n \tag{7.5}$$

is given by (6.7). Moreover, by (7.4) and (7.1),

$$|\Delta \hat{M}_{n,i}| \leq C \sum_{j=0}^{\infty} R_1^{n-\tau_i-j} (N_{i,j} + \mu_j) = C R_1^{n-\tau_i} (\hat{\Xi}_i(R_1^{-1}) + \hat{\mu}(R_1^{-1})). \tag{7.6}$$

Define the random variable  $U := \hat{\Xi}_i(R_1^{-1}) + \hat{\mu}(R_1^{-1})$ . Then  $\mathbb{E} U^2 < \infty$  by (A6), since  $R_1^{-1} < r$ . It follows from (7.6) that, for some  $c > 0$  and every  $\varepsilon > 0$ , defining  $h(x) := \mathbb{E}(U^2 \mathbf{1}\{U > cx\})$ ,

$$\begin{aligned} \mathbb{E}(|\Delta \hat{M}_{n,i}|^2 \mathbf{1}\{|\Delta \hat{M}_{n,i}| > \varepsilon\} \mid \mathcal{G}_{i-1}) &\leq C R_1^{2(n-\tau_i)} \mathbb{E}(U^2 \mathbf{1}\{U > c\varepsilon R_1^{\tau_i-n}\}) \\ &= C R_1^{2(n-\tau_i)} h(\varepsilon R_1^{\tau_i-n}) \\ &\leq C R_1^{2(n-\tau_i)} h(\varepsilon R_1^{-n}). \end{aligned}$$

Thus,

$$\sum_i \mathbb{E}(|\Delta \hat{M}_{n,i}|^2 \mathbf{1}\{|\Delta \hat{M}_{n,i}| > \varepsilon\} \mid \mathcal{G}_{i-1}) \leq C \sum_{k=0}^n B_k R_1^{2(n-k)} h(\varepsilon R_1^{-n}). \tag{7.7}$$

Finally, normalize  $\hat{M}_{n,k}$  by defining  $\tilde{M}_{n,k} := m^{-n/2} \hat{M}_{n,k}$ , thereby yielding a martingale  $(\tilde{M}_{n,k})_k$  with conditional quadratic variation

$$\tilde{V}_n := \sum_i \mathbb{E}((\Delta \tilde{M}_{n,i})^2 \mid \mathcal{G}_{i-1}) = m^{-n} \hat{V}_n \xrightarrow{\text{a.s.}} \sigma^2(\bar{a}) \mathcal{Z}, \tag{7.8}$$

by (7.5), (7.3), and (2.8). Furthermore, by (7.7),

$$\sum_i \mathbb{E}(|\Delta \tilde{M}_{n,i}|^2 \mathbf{1}\{|\Delta \tilde{M}_{n,i}| > \varepsilon\} \mid \mathcal{G}_{i-1}) \leq Ch(\varepsilon m^{n/2} R_1^{-n}) m^{-n} \sum_{k=0}^n B_k R_1^{2(n-k)}, \tag{7.9}$$

which tends to 0 a.s. as  $n \rightarrow \infty$  because  $(m^{1/2} R_1^{-1})^n \rightarrow \infty$ ; hence,  $h(\varepsilon m^{n/2} R_1^{-n}) \rightarrow 0$ , and, by (2.8) and  $R_1^2 < m$ ,

$$m^{-n} \sum_{k=0}^n B_k R_1^{2(n-k)} = m^{-n} \sum_{k=0}^n B_{n-k} R_1^{2k} = \sum_{k=0}^n \frac{B_{n-k}}{m^{n-k}} \left(\frac{R_1^2}{m}\right)^k = O_{\text{a.s.}}(1).$$

The martingales  $(\tilde{M}_{n,i})_i$  thus satisfy a conditional Lindeberg condition, which together with (7.8) implies, by [6, Corollary 3.2], that, as  $n \rightarrow \infty$ , using (7.5),

$$\frac{M_{n,n}}{V_n^{1/2}} = \frac{\hat{M}_{n,Z_n}}{\hat{V}_n^{1/2}} = \frac{\tilde{M}_{n,Z_n}}{\tilde{V}_n^{1/2}} \xrightarrow{D} N(0, 1); \tag{7.10}$$

furthermore, the limit is mixing. (It makes no difference that here we sum the martingale differences to a stopping time  $Z_n$  instead of a deterministic  $k_n$  as in [6].) By (6.5) and (7.3), this yields

$$\frac{\langle \tilde{X}_n, \vec{a} \rangle}{Z_n^{1/2}} \xrightarrow{D} N(0, \sigma^2(\vec{a})). \tag{7.11}$$

We evaluate the asymptotic variance  $\sigma^2(\vec{a})$  given in (7.3) by

$$\begin{aligned} \frac{\sigma^2(\vec{a})}{1 - m^{-1}} &= \sum_{\ell=0}^{\infty} m^{-\ell} \sum_{i,j=1}^{\ell} \sigma_{ij} \alpha_{\ell-i} \alpha_{\ell-j} \\ &= \sum_{k,p=0}^{\infty} \sum_{i,j=1}^{\ell} \sigma_{ij} \alpha_k \alpha_p \mathbf{1}\{i+k = j+p\} m^{-i-k} \\ &= \sum_{k,p,i,j} \sigma_{ij} \alpha_k \alpha_p \oint_{|z|=m^{-1/2}} z^{i+k} \bar{z}^{j+p} \frac{|dz|}{2\pi m^{-1/2}} \\ &= \oint_{|z|=m^{-1/2}} \left| \sum_k \alpha_k z^k \right|^2 \sum_{i,j} \sigma_{ij} z^i \bar{z}^j \frac{|dz|}{2\pi m^{-1/2}}. \end{aligned} \tag{7.12}$$

Furthermore, for  $|z| = m^{-1/2}$  (and any  $z$  with  $|z| < R^{-1} = r$  and  $\hat{\mu}(z) \neq 1$ ), by (6.1) and Lemma 3.6 with  $\lambda = z^{-1}$ ,

$$\sum_{k=0}^{\infty} \alpha_k z^k = \left\langle \sum_{k=0}^{\infty} z^k T^k(\vec{v}), \vec{a} \right\rangle = \langle (1 - zT)^{-1}(\vec{v}), \vec{a} \rangle = \frac{1}{(z-1)(1 - \hat{\mu}(z))} \sum_{\ell} a_{\ell} (z^{\ell} - m^{-\ell}). \tag{7.13}$$

By (7.12)–(7.13),  $\sigma^2(\vec{a})$  equals the right-hand side in (2.15). Thus, (7.11) shows convergence as in (2.14) for any finite linear combination of  $Z_n^{-1/2} X_{n,k}$ , and, thus, joint convergence in (2.14) by the Cramér–Wold device.

Convergence in  $L^2(\ell_R^2)$  follows from this and Lemma 5.1 (with a slightly increased  $R$ ) by a standard truncation argument; we omit the details.

By (2.15), the variable  $\zeta_k$  is degenerate only if  $\Sigma(z) = 0$  for every  $z$  with  $|z| = m^{-1/2}$ , and, thus, by (2.12),  $\hat{\Xi}(z) = \hat{\mu}(z)$  a.s. for every such  $z$ , which, by (2.1)–(2.2), implies that  $N_k = \mu_k$  a.s. for every  $k$ .

### 8. Proof of Theorem 2.2

We assume in this section that  $\gamma_* = m^{-1/2}$  and that (2.16) holds. By Lemma 3.3, the spectral radius  $r(T) = \gamma_*^{-1} = m^{1/2}$ . Lemma 3.5 applies with  $\gamma_* = m^{-1/2}$ , and, thus,  $\tilde{R} < m^{1/2}$ ; we may assume that  $\tilde{R} > R$ .

As in Section 6, fix a real vector  $\vec{a} \in \ell_{R-1}^2$  and define, using (3.31),

$$\beta_i = \beta_i(\vec{a}) := \langle P_i(\vec{v}), \vec{a} \rangle = \langle \vec{v}_i, \vec{a} \rangle = \frac{1}{\gamma_i(\gamma_i - 1)\hat{\mu}'(\gamma_i)} \sum_{k=0}^{\infty} a_k (\gamma_i^k - m^{-k}). \tag{8.1}$$

Then, by (6.1) and Lemma 3.5,

$$\alpha_k = O(\tilde{R}^k) + \sum_{i=1}^q \lambda_i^k \langle P_i(\vec{v}), \vec{a} \rangle = \sum_{i=1}^q \beta_i \lambda_i^k + O(\tilde{R}^k) = O(m^{k/2}). \tag{8.2}$$

Observe that the terms  $O(\cdot)$  in (8.2) hold uniformly in all  $\vec{a}$  for which  $\|\vec{a}\|_{\ell_{R^{-1}}^2} \leq 1$ , as does every term  $O(\cdot)$  in this section. For  $p, t = 1, \dots, q$ , define also

$$\sigma_{pt}^* := \sum_{i,j=1}^{\infty} \sigma_{ij} \lambda_p^{-i} \lambda_t^{-j},$$

and note that, using (6.8),  $|\lambda_p| = m^{1/2}$ , and  $R < m^{1/2}$ ,

$$\sum_{i,j=1}^{\ell} \sigma_{ij} \lambda_p^{-i} \lambda_t^{-j} = \sigma_{pt}^* + O\left(\sum_{i>\ell, j\geq 1} R^{i+j} (m^{1/2})^{-i-j}\right) = \sigma_{pt}^* + O\left(\left(\frac{R}{m^{1/2}}\right)^\ell\right). \tag{8.3}$$

Let

$$s_\ell := \sum_{i,j=1}^{\ell} \sigma_{ij} \alpha_{\ell-i} \alpha_{\ell-j}. \tag{8.4}$$

Then, by (8.2) and symmetry, using (6.8) again,  $|\lambda_p| = m^{1/2}$ , and (8.3),

$$\begin{aligned} s_\ell &:= \sum_{i,j=1}^{\ell} \sigma_{ij} \sum_{p=1}^q \sum_{t=1}^q \beta_p \lambda_p^{\ell-i} \beta_t \lambda_t^{\ell-j} + O\left(\sum_{i,j=1}^{\ell} R^{i+j} m^{(\ell-i)/2} \tilde{R}^{\ell-j}\right) \\ &= \sum_{p=1}^q \sum_{t=1}^q \beta_p \beta_t \lambda_p^\ell \lambda_t^\ell \sum_{i,j=1}^{\ell} \sigma_{ij} \lambda_p^{-i} \lambda_t^{-j} + O((m^{1/2} \tilde{R})^\ell) \\ &= \sum_{p=1}^q \sum_{t=1}^q \beta_p \beta_t \lambda_p^\ell \lambda_t^\ell \sigma_{pt}^* + O((m^{1/2} \tilde{R})^\ell). \end{aligned} \tag{8.5}$$

In particular,

$$s_\ell = O(m^\ell). \tag{8.6}$$

It follows from (2.8), (6.7), (8.4), (8.5), and (8.6) that, a.s.,

$$\begin{aligned} \frac{V_n}{B_n} &= \sum_{\ell=0}^n \frac{B_{n-\ell}}{B_n} s_\ell \\ &= \sum_{\ell=0}^n m^{-\ell} (1 + o(1) + O_{\text{a.s.}}(1) \mathbf{1}\{n - \ell < \log n\}) s_\ell \\ &= \sum_{\ell=0}^n m^{-\ell} s_\ell + o(n) \\ &= \sum_{\ell=0}^n m^{-\ell} \sum_{p=1}^q \sum_{t=1}^q \beta_p \beta_t \lambda_p^\ell \lambda_t^\ell \sigma_{pt}^* + o(n) \\ &= \sum_{p=1}^q \sum_{t=1}^q \beta_p \beta_t \sigma_{pt}^* \sum_{\ell=0}^n \left(\frac{\lambda_p \lambda_t}{m}\right)^\ell + o(n). \end{aligned} \tag{8.7}$$

Recall that  $|\lambda_p| = |\lambda_t| = m^{1/2}$ , so  $|\lambda_p \lambda_t / m| = 1$ , and if  $\lambda_t = \bar{\lambda}_p$  then  $\sum_{\ell=0}^n (\lambda_p \lambda_t / m)^\ell = n + 1$ , while if  $\lambda_t \neq \bar{\lambda}_p$  then  $\sum_{\ell=0}^n (\lambda_p \lambda_t / m)^\ell = O(1)$ . Then, since by (3.1) and (2.8),  $B_n / Z_n \xrightarrow{\text{a.s.}} 1 - m^{-1}$ , (8.7) yields

$$\begin{aligned} \frac{V_n}{nZ_n} &\xrightarrow{\text{a.s.}} \sigma^2(\vec{a}) \\ &:= \frac{m-1}{m} \sum_{p=1}^q \sum_{t=1}^q \beta_p \beta_t \sigma_{pt}^* \mathbf{1}\{\lambda_t = \bar{\lambda}_p\} \\ &= \frac{m-1}{m} \sum_{p=1}^q |\beta_p|^2 \sum_{i,j=1}^\infty \sigma_{ij} \lambda_p^{-i} \bar{\lambda}_p^{-j} \\ &= \frac{m-1}{m} \sum_{p=1}^q |\beta_p|^2 \Sigma(\gamma_p). \end{aligned} \tag{8.8}$$

We refine the martingale  $(M_{n,k})_k$  to  $(\hat{M}_{n,k})_k$  as in Section 7, normalizing it now to  $\tilde{M}_{n,k} := (nm^n)^{-1/2} \hat{M}_{n,k}$ . It follows from (2.8) and (8.8) that the conditional quadratic variation  $\tilde{V}_n = V_n / (nm^n) \xrightarrow{\text{a.s.}} \sigma^2(\vec{a})Z$ , i.e. (7.8) holds in the present case as well. Furthermore, if we now let  $R_1 := m^{1/2}$ , then (7.1) and (7.6)–(7.7) hold, and it follows that (7.9) is modified to

$$\begin{aligned} \sum_i \mathbb{E}(|\Delta \tilde{M}_{n,i}|^2 \mathbf{1}\{|\Delta \tilde{M}_{n,i}| > \varepsilon\} \mid \mathcal{G}_{i-1}) &\leq Ch(\varepsilon n^{1/2}) \frac{1}{nm^n} \sum_{k=0}^n B_k m^{n-k} \\ &= O_{\text{a.s.}}(h(\varepsilon n^{1/2})) \\ &\xrightarrow{\text{a.s.}} 0. \end{aligned}$$

Hence, the conditional Lindeberg condition holds in the present case too, and (7.10) again follows by [6, Corollary 3.2], which now, by (8.8) and (6.5), yields (mixing)

$$\frac{\langle \vec{X}_n, \vec{a} \rangle}{(nZ_n)^{1/2}} \xrightarrow{D} N(0, \sigma^2(\vec{a})).$$

By (8.8) and (8.1), this proves (2.17)–(2.18).

By (2.18), the variable  $\zeta_k$  is degenerate only if  $\Sigma(\gamma_p) = 0$  for every  $p$ , and, thus, by (2.12),  $\hat{\mathbb{E}}(\gamma_p) = \hat{\mu}(\gamma_p)$  a.s.

As in Section 7, convergence in  $L^2(\ell_R^2)$  follows by a standard truncation argument, using now Lemma 8.1 below (with an increased  $R$ ); we omit the details.

**Lemma 8.1.** *Assume that (A1)–(A6) hold,  $\gamma_* = m^{-1/2}$ , and (2.16) holds. If  $R < m^{1/2}$  then, for all  $n, k \geq 0$ ,*

$$\mathbb{E} X_{n,k}^2 \leq Cnm^n R^{-2k} \tag{8.9}$$

and

$$\mathbb{E} \|\vec{X}_n\|_{\ell_R^2}^2 \leq Cnm^n.$$

*Proof.* By (2.7), (6.5), (6.7), (8.4), and (8.6),

$$\mathbb{E} \langle \vec{X}_n, \vec{a} \rangle^2 = \mathbb{E} V_n = \mathbb{E} \sum_{\ell=0}^n B_{n-\ell} s_\ell \leq Cnm^n,$$

uniformly for  $\|\vec{a}\|_{\ell_{R^{-1}}}^2 \leq 1$ . Setting  $\vec{a} = R^k(\delta_{kj})_j$  yields (8.9). Applying (8.9) with  $R$  replaced

by some  $R'$  for which  $R < R' < m^{1/2}$ ,

$$\mathbb{E} \|\vec{X}_n\|_{\ell_R^2}^2 = \sum_{k=0}^{\infty} R^{2k} \mathbb{E} X_{n,k}^2 \leq Cnm^n \sum_{k=0}^{\infty} \left(\frac{R}{R'}\right)^{2k} = Cnm^n. \quad \square$$

**9. Proof of Theorem 2.3**

Assume now that  $\gamma_* < m^{-1/2}$ . By Lemma 3.3, the spectral radius  $r(T) = \gamma_*^{-1} \geq m^{1/2}$ . We apply Lemma 3.5, assuming as we may that  $\tilde{R} > m^{1/2}$ . (Otherwise we increase  $\tilde{R}$ , keeping  $\tilde{R} < \gamma_*^{-1}$ .) Hence, by (3.29),

$$T^k(\vec{v}) = T_0^k(\vec{v}) + \sum_{i=1}^q \lambda_i^k P_i(\vec{v}) = T_0^k(\vec{v}) + \sum_{i=1}^q \lambda_i^k \vec{v}_i.$$

Then, by (3.17),

$$\vec{X}_n = - \sum_{k=0}^n W_k (T P_0)^{n-k}(\vec{v}) - \sum_{i=1}^q \sum_{k=0}^n \lambda_i^{n-k} W_k \vec{v}_i. \tag{9.1}$$

Recall (3.8) and let

$$\tilde{U}_i := - \sum_{k=0}^{\infty} \gamma_i^k W_k = - \sum_{k=0}^{\infty} \lambda_i^{-k} W_k = - \sum_{\ell=0}^{\infty} \sum_{j=1}^{\infty} \lambda_i^{-\ell-j} W_{\ell,j}, \tag{9.2}$$

noting that, by Lemma 3.1 and  $|\gamma_i| = \gamma_* < m^{-1/2}$ , the sum converges in  $L^2$  and

$$\left\| \tilde{U}_i + \sum_{k=0}^n \lambda_i^{-k} W_k \right\|_2 \leq \sum_{k=n+1}^{\infty} C |\lambda_i|^{-k} m^{k/2} \leq C(\gamma_* m^{1/2})^n. \tag{9.3}$$

Furthermore, by Lemma 3.1 and (3.30), since  $\tilde{R} > m^{1/2}$ ,

$$\left\| \sum_{k=0}^n W_k (T P_0)^{n-k}(\vec{v}) \right\|_{L^2(\ell_R^2)} \leq \sum_{k=0}^n \|W_k\|_2 \|(T P_0)^{n-k}(\vec{v})\|_{\ell_R^2} \leq C \sum_{k=0}^n m^{k/2} \tilde{R}^{n-k} \leq C \tilde{R}^n. \tag{9.4}$$

By (9.1), (9.3), and (9.4), defining  $U_i := (\gamma_i(\gamma_i - 1)\hat{\mu}'(\gamma_i))^{-1} \tilde{U}_i$  so, by (3.31),  $\tilde{U}_i \vec{v}_i = U_i \vec{u}_i$ ,

$$\begin{aligned} \left\| \gamma_*^n \vec{X}_n - \sum_{i=1}^q \left(\frac{\lambda_i}{|\lambda_i|}\right)^n U_i \vec{u}_i \right\|_{L^2(\ell_R^2)} &\leq C \gamma_*^n \tilde{R}^n + \sum_{i=1}^q \left\| \sum_{k=0}^n \lambda_i^{-k} W_k \vec{v}_i + \tilde{U}_i \vec{v}_i \right\|_{L^2(\ell_R^2)} \\ &\leq C(\gamma_* \tilde{R})^n + C(\gamma_* m^{1/2})^n \\ &\leq C(\gamma_* \tilde{R})^n. \end{aligned} \tag{9.5}$$

Since  $\gamma_* \tilde{R} < 1$ , this shows convergence in  $L^2(\ell_R^2)$  in (2.19), and convergence a.s. follows by (9.5) and the Borel–Cantelli lemma.

We have  $\mathbb{E} U_i = \mathbb{E} \tilde{U}_i = 0$  by (9.2) since  $\mathbb{E} W_k = 0$  by (3.6)–(3.8). Furthermore,  $W_{0,k} = B_{0,k} - \mu_k = N_k - \mu_k$ , while  $\mathbb{E}(W_{n,k} | \mathcal{F}_0) = 0$  for  $n \geq 1$  by (3.6); hence by (3.8),  $\mathbb{E}(W_n | \mathcal{F}_0) = W_{0,n} = N_n - \mu_n$ , and, thus,

$$\mathbb{E}(\tilde{U}_i | \mathcal{F}_0) = - \sum_{k=0}^{\infty} \gamma_i^k (N_k - \mu_k) = -\hat{\Xi}(\gamma_i) + \hat{\mu}(\gamma_i).$$

Thus,  $U_i$  is degenerate only if  $\hat{\Xi}(\gamma_i)$  is so.



### 10. A stochastic integral calculus

The limit variables  $\zeta_k$  in Theorems 2.1 and 2.2 can be interpreted as stochastic integrals of certain functions (‘symbols’). This interpretation leads to a useful symbolic calculus. There are also some partial related results for Theorem 2.3. We consider the three cases in Theorems 2.1–2.3 separately.

#### 10.1. The case $\gamma_* > m^{-1/2}$

Assume throughout this subsection that Theorem 2.1 applies; in particular that  $\gamma_* > m^{-1/2}$ . Let  $\nu$  be the finite measure on the circle  $|z| = m^{-1/2}$  given by

$$d\nu(z) := \frac{m-1}{m} |1-z|^{-2} |1-\hat{\mu}(z)|^{-2} \Sigma(z) \frac{|dz|}{2\pi m^{-1/2}}, \tag{10.1}$$

and consider an isomorphism  $\mathfrak{I} : L^2(\nu) \rightarrow \mathcal{H}$  of the Hilbert space  $L^2(\nu)$  into a Gaussian Hilbert space  $\mathcal{H}$ , i.e. a Hilbert space of Gaussian random variables;  $\mathfrak{I}$  can be interpreted as a stochastic integral (see [8, Section VII.2]). Here we let  $L^2(\nu)$  be the space of complex square-integrable functions, but regard it as a real Hilbert space with the inner product  $\langle f, g \rangle_\nu := \text{Re} \int f \bar{g} d\nu$ . Then (2.14)–(2.15) can be restated as

$$Z_n^{-1/2} X_{n,k} \xrightarrow{D} \zeta_k := \mathfrak{I}(z^k - m^{-k}) \quad \text{as } n \rightarrow \infty, \text{ jointly for all } k \geq 0. \tag{10.2}$$

This yields a convenient calculus for joint limits as the next three examples illustrate.

**Example 10.1.** Let  $k, \ell \geq 0$ . Then, by (2.10),

$$X_{n-\ell,k} = X_{n,k+\ell} - m^{-k} X_{n,\ell},$$

and, thus, recalling (2.9), jointly for all  $k, \ell \geq 0$ ,

$$\begin{aligned} Z_{n-\ell}^{-1/2} X_{n-\ell,k} &\xrightarrow{D} m^{\ell/2} (\zeta_{k+\ell} - m^{-k} \zeta_\ell) \\ &= m^{\ell/2} \mathfrak{I}(z^{k+\ell} - m^{-k} z^\ell) \\ &= \mathfrak{I}((zm^{1/2})^\ell (z^k - m^{-k})). \end{aligned} \tag{10.3}$$

Denoting this limit by  $\zeta_k^{(\ell)}$ , we have of course  $\zeta_k^{(\ell)} \stackrel{D}{=} \zeta_k$ , which corresponds to the fact that  $|zm^{1/2}|^\ell = 1$  on the support of  $\nu$ . More interesting is the joint convergence

$$(Z_n^{-1/2} X_{n,k}, Z_{n-\ell}^{-1/2} X_{n-\ell,k}) \xrightarrow{D} (\zeta_k, \zeta_k^{(\ell)}),$$

with covariance

$$\begin{aligned} \text{cov}(\zeta_k, \zeta_k^{(\ell)}) &= \langle z^k - m^{-k}, (zm^{1/2})^\ell (z^k - m^{-k}) \rangle_\nu \\ &= \text{Re} \int_{|z|=m^{-1/2}} (zm^{1/2})^\ell |z^k - m^{-k}|^2 d\nu. \end{aligned} \tag{10.4}$$

Equation (10.1) shows that the measure  $\nu$  is absolutely continuous on the circle  $|z| = m^{-1/2}$ . The change of variables  $z = m^{-1/2} e^{i\theta}$  yields  $(zm^{1/2})^\ell = e^{i\ell\theta}$  and the Riemann–Lebesgue lemma shows that, for every fixed  $k$ ,  $\text{cov}(\zeta_k, \zeta_k^{(\ell)}) \rightarrow 0$  as  $\ell \rightarrow \infty$ . Roughly speaking,  $X_{n-\ell,k}$  and  $X_{n,k}$  are thus essentially uncorrelated when  $\ell$  is large, which justifies the claim in Section 2 that there is only short-range dependence in this case.

**Example 10.2.** We can define  $X_{n,k}$  by (2.10) also for  $k < 0$ . Then the calculations in Example 10.1 apply to any  $\ell \geq 0$  and any  $k \geq -\ell$ . In (10.3) replace  $n$  by  $n + \ell$ , so, for any fixed  $\ell$ , as  $n \rightarrow \infty$ ,

$$Z_n^{-1/2} X_{n,k} \xrightarrow{D} \mathfrak{L}((zm^{1/2})^\ell (z^k - m^{-k})),$$

jointly for all  $k \geq -\ell$ . Since the factor  $(zm^{1/2})^\ell$  does not depend on  $k$  and has absolute value 1, this means (by changing the isomorphism  $\mathfrak{L}$ ) that (10.2) holds jointly for all  $k \geq -\ell$ . Since  $\ell$  is arbitrary, this means that (10.2) holds jointly for all  $k \in \mathbb{Z}$ . Hence, (2.14)–(2.15) extend to all  $k \in \mathbb{Z}$ , as claimed in Remark 2.1.

**Example 10.3.** From (2.10),

$$m^{-j} Z_{n+j} - m^{-j-1} Z_{n+j+1} = m^{-j} X_{n+j+1,1}.$$

Hence, by Lemma 5.1, for  $j \geq 0$ ,

$$\|m^{-j} Z_{n+j} - m^{-j-1} Z_{n+j+1}\|_2 \leq C m^{-j+(n+j+1)/2} = C m^{n/2-j/2}. \tag{10.5}$$

Summing (10.5) for  $j \geq \ell$  we obtain, recalling (2.8),

$$\|m^{-\ell} Z_{n+\ell} - m^n \mathfrak{Z}\|_2 \leq C m^{n/2-\ell/2}$$

for  $n \geq 1$  and  $\ell \geq 0$ . Hence, as  $\ell \rightarrow \infty$ ,  $m^{-n/2}(m^{-\ell} Z_{n+\ell} - m^n \mathfrak{Z}) \rightarrow 0$  in  $L^2$ , and, thus, in probability, uniformly in  $n$ . Since  $Z_n/m^n \xrightarrow{\text{a.s.}} \mathfrak{Z} > 0$ , and, thus,  $\sup_n m^n/Z_n < \infty$  a.s., it follows that, still uniformly in  $n$ ,

$$Z_n^{-1/2} (m^{-\ell} Z_{n+\ell} - m^n \mathfrak{Z}) \xrightarrow{P} 0 \quad \text{as } \ell \rightarrow \infty. \tag{10.6}$$

Define the random variables

$$Y_{n,\ell} := Z_n^{-1/2} (Z_n - m^{-\ell} Z_{n+\ell}) = -Z_n^{-1/2} m^{-\ell} X_{n,-\ell}, \quad \ell \geq 0.$$

Then, by (10.2) and Example 10.2, for every fixed  $\ell$ ,

$$Y_{n,\ell} \xrightarrow{D} -m^{-\ell} \zeta_{-\ell} = \mathfrak{L}(1 - m^{-\ell} z^{-\ell}) \quad \text{as } n \rightarrow \infty. \tag{10.7}$$

Also, by (10.6),  $Y_{n,\ell} \xrightarrow{P} Z_n^{-1/2} (Z_n - m^n \mathfrak{Z})$  as  $\ell \rightarrow \infty$ , uniformly in  $n$ . Finally,  $|mz| = m^{1/2} > 1$  on the support of  $\nu$ , and, thus,  $1 - (mz)^{-\ell} \rightarrow 1$  in  $L^2(\nu)$  as  $\ell \rightarrow \infty$ ; hence,  $\mathfrak{L}(1 - m^{-\ell} z^{-\ell}) \rightarrow \mathfrak{L}(1)$  as  $\ell \rightarrow \infty$ , in  $L^2$  and, thus, in distribution. It follows that we can let  $\ell \rightarrow \infty$  in (10.7) (see [2, Theorem 4.2]), leading to

$$Z_n^{-1/2} (Z_n - m^n \mathfrak{Z}) \xrightarrow{D} \mathfrak{L}(1) \quad \text{as } n \rightarrow \infty.$$

The convergence here and for all  $k \geq 0$  in (10.2) holds jointly. Hence, jointly for all  $k \in \mathbb{Z}$ ,

$$Z_n^{-1/2} (Z_{n-k} - m^{n-k} \mathfrak{Z}) = Z_n^{-1/2} (X_{n,k} + m^{-k} (Z_n - m^n \mathfrak{Z})) \xrightarrow{D} \mathfrak{L}(z^k). \tag{10.8}$$

Conversely, (10.2) follows immediately from (10.8).

In the Galton–Watson case (Example 2.1), (10.8) is equivalent to the case  $q = 0$  of [7, Theorem (2.10.2)].

**10.2. The case  $\gamma_* = m^{-1/2}$**

Assume now that Theorem 2.2 applies; thus,  $\gamma_* = m^{-1/2}$  and (2.16) holds. Here let  $\nu$  be the discrete measure, with support  $\Gamma_{**}$ ,

$$\nu(\cdot) := (m - 1) \sum_{p=1}^q |1 - \gamma_p|^{-2} |\hat{\mu}'(\gamma_p)|^{-2} \Sigma(\gamma_p) \delta_{\gamma_p}(\cdot),$$

and consider an isomorphism  $I$  of  $L^2(\nu)$  into a Gaussian Hilbert space as above. Then (2.17)–(2.18) can be stated as (10.2), with the normalizing factor changed from  $Z_n^{-1/2}$  to  $(nZ_n)^{-1/2}$ . With this changed normalization of  $X_{n,k}$ , all results in the preceding subsection hold, with one exception: the measure  $\nu$  now has fixed finite support, so there exists a sequence  $\ell_j \rightarrow \infty$  such that  $(zm^{1/2})^{\ell_j} \rightarrow 1$  as  $j \rightarrow \infty$  for every  $z \in \text{supp}(\nu) = \Gamma_{**}$ ; hence, (10.4) implies that  $\limsup_{\ell \rightarrow \infty} \text{corr}(\zeta_k, \zeta_k^{(\ell)}) = 1$ . Hence, while the convergence in (2.18) is mixing so there is no dependence on the initial generations as in the case  $\gamma_* < m^{-1/2}$ , there does exist dependence over a larger range than in the case  $\gamma_* > m^{-1/2}$ .

Moreover, each  $\zeta_k$  now belongs to the space spanned by  $\zeta_1, \dots, \zeta_q$ , typically  $q$ -dimensional; this yields the linear dependence of the limits  $\zeta_k$  claimed in Section 2.

**Example 10.4.** In the simplest case,  $\Gamma_{**} = \{-m^{1/2}\}$  (see Example 2.2 for example). Then  $\zeta_k = ((-1)^k m^{-k/2} - m^{-k})\zeta$  for some  $\zeta \sim N(0, \nu\{-m^{1/2}\})$  and all  $k \in \mathbb{Z}$ .

Furthermore,  $zm^{1/2} = -1$  on  $\text{supp } \nu$ , and, thus, (10.3) yields  $\zeta_k^{(\ell)} = (-1)^\ell \zeta_k$ ; in particular,  $\zeta_k^{(\ell)} = \zeta_k$  for every even  $\ell$ .

**10.3. The case  $\gamma_* < m^{-1/2}$**

In this case, there is no limit, but we can argue with the components of the approximating sum in (2.19) in the same way as with  $\zeta_k$  in Examples 10.1–10.2, and conclude that (2.19), interpreted componentwise, extends also to  $k < 0$ , as claimed in Remark 2.1. We omit details.

**11. Random characteristics**

A random characteristic is a random function  $\chi(t): [0, \infty) \rightarrow \mathbb{R}$  defined on the same probability space as the prototype offspring process  $\Xi$  (with  $\chi(t) \equiv 0$  for  $t < 0$ ); assume that each individual  $x$  has an independent copy  $(\Xi_x, \chi_x)$  of  $(\Xi, \chi)$ , and interpret  $\chi_x(t)$  as the characteristic of  $x$  at age  $t$ . Consider the lattice case as above. Denoting the birth time of  $x$  by  $\tau_x$ , define

$$Z_n^\chi := \sum_{\{x: \tau_x \leq n\}} \chi_x(n - \tau_x), \tag{11.1}$$

the total of the characteristic over all individuals at time  $n$  (see, e.g. [7, Chapter 6] for more detail). We make the following assumption.

(C) There exists  $R_2 < m^{1/2}$  such that  $\mathbb{E}[\chi(k)^2] \leq CR_2^{2k}$  for some  $C < \infty$  and all  $k \geq 0$ .

Define

$$\lambda_k^\chi := \begin{cases} \mathbb{E} \chi(k) & \text{for } k \geq 0, \\ 0 & \text{otherwise,} \end{cases}$$

$$\Lambda^\chi(z) := \sum_{k=0}^{\infty} \lambda_k^\chi z^k, \tag{11.2}$$

$$\lambda^\chi := (1 - m^{-1})\Delta^\chi(m^{-1}) = \sum_{k=0}^\infty (m^{-k} - m^{-k-1})\lambda_k^\chi, \tag{11.3}$$

$$\kappa_{j,k} := \text{cov}(\chi(j), N_k).$$

Note that (C) implies that

$$|\lambda_k^\chi| = |\mathbb{E} \chi(k)| \leq CR_2^k, \tag{11.4}$$

implying in turn that the sum in (11.2) converges absolutely at least for  $|z| \leq m^{-1/2}$ , and, hence, the sum in (11.3) converges absolutely.

Decompose the characteristic into its mean  $\lambda_k^\chi = \mathbb{E} \chi(k)$  and the centered part

$$\tilde{\chi}(k) := \chi(k) - \mathbb{E} \chi(k) = \chi(k) - \lambda_k^\chi,$$

and define

$$V_{n,k}^\chi := \sum_{\{x: \tau_x=n\}} \tilde{\chi}_x(k) = \sum_{\{x: \tau_x=n\}} (\chi_x(k) - \lambda_k^\chi) = \sum_{\{x: \tau_x=n\}} \chi_x(k) - \lambda_k^\chi B_n.$$

Then (11.1) implies that

$$Z_n^{\tilde{\chi}} = \sum_{k=0}^n V_{n-k,k}^\chi = \sum_{k=0}^\infty V_{n-k,k}^\chi, \quad Z_n^\chi = \sum_{k=0}^n (V_{n-k,k}^\chi + \lambda_k^\chi B_{n-k}) = Z_n^{\tilde{\chi}} + \sum_{k=0}^\infty \lambda_k^\chi B_{n-k}.$$

Recalling (2.10), (3.1), and (11.3), leads to the decomposition

$$\begin{aligned} Z_n^\chi - \lambda^\chi Z_n &= Z_n^{\tilde{\chi}} + \sum_{k=0}^\infty \lambda_k^\chi (B_{n-k} - (m^{-k} - m^{-k-1})Z_n) \\ &= Z_n^{\tilde{\chi}} + \sum_{k=0}^\infty \lambda_k^\chi (X_{n,k} - X_{n,k+1}) \\ &= Z_n^{\tilde{\chi}} + \sum_{k=1}^n (\lambda_k^\chi - \lambda_{k-1}^\chi) X_{n,k} \\ &= Z_n^{\tilde{\chi}} + \langle \vec{X}_n, \Delta \vec{\lambda}^\chi \rangle, \end{aligned} \tag{11.5}$$

where  $\Delta \vec{\lambda}^\chi$  is the vector  $(\lambda_k^\chi - \lambda_{k-1}^\chi)_{k=0}^\infty$ . Here  $\Delta \vec{\lambda}^\chi \in \ell^2_{R^{-1}}$  by (11.4), and, thus, the asymptotic behaviour of  $\langle \vec{X}_n, \Delta \vec{\lambda}^\chi \rangle$  is given by Theorems 2.1–2.3.

The term  $Z_n^{\tilde{\chi}}$  in (11.5) is asymptotically normal after normalization, for any value of  $\gamma_*$ , as shown in the next theorem in which the assumption  $\mathbb{E} \chi(k) = 0$  is equivalent to  $\chi = \tilde{\chi}$ .

**Theorem 11.1.** *Assume that (A1)–(A6) and (C) hold. If  $\mathbb{E} \chi(k) = 0$  for every  $k \geq 0$  then, as  $n \rightarrow \infty$ ,*

$$Z_n^{-1/2} Z^\chi \xrightarrow{D} \zeta^\chi \tag{11.6}$$

for some normal random variable  $\zeta^\chi$  with mean  $\mathbb{E} \zeta^\chi = 0$  and variance

$$\text{var}(\zeta^\chi) = \frac{m-1}{m} \sum_{k=0}^\infty m^{-k} \text{var}(\chi(k)). \tag{11.7}$$

Before proving Theorem 11.1, note that, when  $\gamma_* > m^{-1/2}$ , Theorems 11.1 and 2.1 show that in (11.5) each of  $Z_n^\chi$  and  $\langle \bar{X}_n, \Delta \bar{\lambda}^\chi \rangle$  is asymptotically normal after normalization by  $Z_n^{1/2}$ . Indeed, as shown below, in this case the two terms are jointly asymptotically normal, leading by (11.5) to the following extension of Theorem 2.1 which is the deterministic case  $\chi(k) = \sum_{j \leq k} a_j$ .

**Theorem 11.2.** *Assume that (A1)–(A6), (B), and (C) hold. Then, as  $n \rightarrow \infty$ ,*

$$Z_n^{-1/2} (Z^\chi - \lambda^\chi Z_n) \xrightarrow{D} \zeta^\chi$$

for some normal random variable  $\zeta^\chi$  with mean  $\mathbb{E} \zeta^\chi = 0$  and variance as in

$$\begin{aligned} \frac{m \operatorname{var}(\zeta^\chi)}{m-1} &= \sum_{k=0}^\infty m^{-k} \operatorname{var}(\chi(k)) \\ &- 2 \oint_{|z|=m^{-1/2}} \frac{(1-z)\Lambda^\chi(z) - \lambda^\chi}{(z-1)(1-\hat{\mu}(z))} \sum_{k=0}^\infty \sum_{j=1}^\infty \kappa_{kj} z^j \bar{z}^k \frac{|dz|}{2\pi m^{-1/2}} \\ &+ \oint_{|z|=m^{-1/2}} \frac{|(1-z)\Lambda^\chi(z) - \lambda^\chi|^2}{|1-z|^2 |1-\hat{\mu}(z)|^2} \sum_{i,j} \sigma_{ij} z^i \bar{z}^j \frac{|dz|}{2\pi m^{-1/2}}. \end{aligned} \tag{11.8}$$

**Remark 11.1.** In both Theorems 11.1 and 11.2, joint asymptotic normality for several characteristics, with a corresponding formula for asymptotic covariances, follow by the proof, or by the Cramér–Wold device.

*Proofs of Theorems 11.1 and 11.2.* We use results from Section 5, and assume as we may that  $R$  is so chosen that  $R_2 < R < m^{1/2}$ .

Given  $B_{n-k}$ ,  $V_{n-k,k}^\chi$  is the sum of  $B_{n-k}$  independent copies of  $\tilde{\chi}(k) = \chi(k) - \mathbb{E} \chi(k)$ . Then using (C), (2.7), and  $B_{n-k} \leq Z_{n-k}$ ,

$$\mathbb{E}[(V_{n-k,k}^\chi)^2] = \mathbb{E}[\mathbb{E} (V_{n-k,k}^\chi)^2 \mid B_{n-k}] = \operatorname{var}(\chi(k)) \mathbb{E} B_{n-k} \leq C m^{n-k} R_2^{2k},$$

and, using (11.4) and Lemma 5.1,

$$\mathbb{E}[(\lambda_k^\chi (X_{n,k} - X_{n,k+1}))^2] \leq C R_2^{2k} (\mathbb{E} X_{n,k}^2 + \mathbb{E} X_{n,k+1}^2) \leq C m^n \left(\frac{R_2}{R}\right)^{2k}.$$

Because  $R_2 < R < m^{1/2}$ , it follows by standard arguments that on replacing  $\chi$  by the truncated characteristic  $\chi_K(k) := \chi(k) \mathbf{1}\{k \leq K\}$ , the error  $Z_n^{-1/2} (Z_n^\chi - \lambda^\chi Z_n - (Z_n^{\chi_K} - \lambda^{\chi_K} Z_n))$  tends to 0 in probability as  $K \rightarrow \infty$ , uniformly in  $n$ , and, consequently (see [2, Theorem 4.2]), it suffices to prove both theorems for the truncated characteristic  $\chi_K$ . Hence, in the sequel we can assume, with a change in notation, that, for some  $K < \infty$ ,  $\chi(k) = 0$  for  $k > K$ .

Let  $\vec{\vartheta} = (\vartheta_0, \vartheta_1, \dots)$  be a random vector such that  $(\vec{\vartheta}, \vec{\eta})$  is jointly normal with mean 0 and covariances given by (4.1) and

$$\begin{aligned} \operatorname{cov}(\vartheta_j, \vartheta_k) &= \operatorname{cov}(\chi(j), \chi(k)), \\ \operatorname{cov}(\vartheta_j, \eta_k) &= \kappa_{j,k} := \operatorname{cov}(\chi(j), N_k). \end{aligned} \tag{11.9}$$

Let  $(\vec{\vartheta}^{(k)}, \vec{\eta}^{(k)})$  be independent copies of  $(\vec{\vartheta}, \vec{\eta})$ . The proof of Lemma 4.1 extends to show that (4.2) holds jointly with

$$Z_n^{-1/2} V_{n-k,k}^\chi \xrightarrow{D} (1 - m^{-1})^{1/2} m^{-k/2} \vartheta_k^{(k)}, \quad k \geq 0. \tag{11.10}$$

Summing (11.10) over  $k \leq K$ , we obtain

$$Z_n^{-1/2} Z_n^{\tilde{\chi}} \xrightarrow{D} \zeta^\chi := (1 - m^{-1})^{1/2} \sum_{k=0}^\infty m^{-k/2} \vartheta_k^{(k)}, \tag{11.11}$$

which yields (11.6) and (11.7) in the case  $\chi = \tilde{\chi}$ ; recall that the terms  $\vartheta_k^{(k)}$  are independent. This completes the proof of Theorem 11.1.

It remains to prove Theorem 11.2, so assume that (B) holds. We have just shown that (4.2) holds jointly with (11.10). Hence, by the proof in Section 5, (5.8) holds jointly with (11.10) for all  $k$ , and, thus, also with (11.11). Consequently, by (11.5),

$$\begin{aligned} & (1 - m^{-1})^{-1/2} Z_n^{-1/2} (Z_n^\chi - \lambda^\chi Z_n) \\ & \xrightarrow{D} \sum_{k=0}^\infty m^{-k/2} \vartheta_k^{(k)} - \sum_{k=0}^\infty \sum_{j=1}^\infty m^{-(k+j)/2} \eta_j^{(k+j)} \langle T^k(\vec{v}), \Delta \vec{\lambda}^\chi \rangle \\ & =: A_1 - A_2, \quad \text{say.} \end{aligned}$$

Here,  $A_1$  and  $A_2$  are jointly normal with means 0. It remains to calculate  $\text{var}(A_1 - A_2)$ . The terms in the sum defining  $A_1$  are independent, so (cf. (11.7) and (11.11))

$$\text{var}(A_1) = \sum_{k=0}^\infty m^{-k} \text{var}(\vartheta_k) = \sum_{k=0}^\infty m^{-k} \text{var}(\chi(k)); \tag{11.12}$$

this yields the first term in (11.8). In Section 5 we calculated  $\text{var}(A_2)$  (see (2.15) and (5.10)), and this yields the last term in (11.8), using  $\sum_k (\lambda_k^\chi - \lambda_{k-1}^\chi) z^k = (1 - z)\Lambda^\chi(z)$  and (11.3). Finally, using (11.9) and (5.12),

$$\begin{aligned} \text{cov}(A_1, A_2) &= \sum_{k=0}^\infty \sum_{j=1}^\infty m^{-(k+j)} \kappa_{k+j,j} \langle T^k(\vec{v}), \Delta \vec{\lambda}^\chi \rangle \\ &= \sum_{k=0}^\infty \sum_{j=1}^\infty \langle T^k(\vec{v}), \Delta \vec{\lambda}^\chi \rangle \oint_{|z|=m^{-1/2}} z^{k+j} \sum_{\ell=0}^\infty \bar{z}^\ell \kappa_{\ell,j} \frac{|dz|}{2\pi m^{-1/2}} \\ &= \oint_{|z|=m^{-1/2}} \langle (1 - zT)^{-1}(\vec{v}), \Delta \vec{\lambda}^\chi \rangle \sum_{\ell=0}^\infty \sum_{j=1}^\infty z^j \bar{z}^\ell \kappa_{\ell,j} \frac{|dz|}{2\pi m^{-1/2}} \\ &= \oint_{|z|=m^{-1/2}} \frac{(1 - z)\Lambda^\chi(z) - (1 - m^{-1})\Lambda^\chi(m^{-1})}{(z - 1)(1 - \hat{\mu}(z))} \\ & \quad \times \sum_{\ell=0}^\infty \sum_{j=1}^\infty z^j \bar{z}^\ell \kappa_{\ell,j} \frac{|dz|}{2\pi m^{-1/2}}. \end{aligned} \tag{11.13}$$

Combining (2.15), (11.12), and (11.13), and recalling (11.3), result (11.8) follows. □

Theorem 11.2 yields asymptotic normality of  $Z_n^\chi$  when  $\gamma_* > m^{-1/2}$ , and Theorem 11.1 shows the same for any  $\gamma_*$  in the special case that  $\mathbb{E} \chi(k) = 0$  for every  $k$ . It remains to consider the case that  $\lambda_k^\chi = \mathbb{E} \chi(k) \neq 0$  for some  $k$  and  $\gamma_* \leq m^{-1/2}$ . If  $\gamma_* = m^{-1/2}$  and (2.16) holds, then Theorem 2.2 shows that  $\langle \bar{X}_n, \Delta \vec{\lambda}^\chi \rangle / \sqrt{n Z_n} \xrightarrow{D} N(0, \sigma^2)$ , where  $\sigma^2$  is given by (2.18) and  $\sigma^2 > 0$

except in degenerate cases. Since Theorem 11.1 implies that  $Z_n^{\tilde{\chi}}/\sqrt{nZ_n} \xrightarrow{\mathbb{P}} 0$ , it follows from (11.5) that  $(Z_n^{\tilde{\chi}} - \lambda^{\tilde{\chi}}Z_n)/\sqrt{nZ_n} \xrightarrow{\mathbb{D}} N(0, \sigma^2)$ . Similarly, if  $\gamma_* < m^{-1/2}$  then Theorem 11.1 implies that  $\gamma_*^n Z_n^{\tilde{\chi}} \xrightarrow{\mathbb{P}} 0$ , and (11.5) shows that  $Z_n^{\tilde{\chi}} - \lambda^{\tilde{\chi}}Z_n$  has the same oscillatory asymptotic behaviour as  $(\tilde{X}_n, \Delta\tilde{\lambda}^{\tilde{\chi}})$ , given by Theorem 2.3.

Summarizing, if  $\gamma_* \leq m^{-1/2}$  then the randomness in the characteristic  $\chi$  only gives an effect of smaller order than the mean  $\mathbb{E}\chi$ , and unless the mean vanishes (or the limits degenerate),  $Z_n^{\tilde{\chi}}$  has the same asymptotic behaviour as if  $\chi$  is replaced by the deterministic  $\mathbb{E}\chi$ , which is treated by Theorems 2.2 and 2.3.

**Example 11.1.** In this paper we have for simplicity assumed (A4), that there are no deaths. Suppose now, more generally, that each individual has a random lifelength  $\ell \leq \infty$ , as usual with i.i.d. copies  $(\Xi_x, \ell_x)$  for all individuals  $x$ . The results in Section 2 apply if we ignore deaths and let  $Z_n$  denote the number of individuals born up to time  $n$ , living or dead. Moreover, the number of living individuals at time  $n$  is  $Z_n^{\chi}$ , for the characteristic  $\chi(k) := \mathbf{1}\{\ell > k\}$ .

Similarly, for example, the number of individuals living at time  $n - j$  is  $Z_n^{\chi_j}$  with

$$\chi_j(k) := \mathbf{1}\{\ell > k - j \geq 0\}.$$

The analogue of  $X_{n,j}$  in (2.10) but counting only living individuals is thus given by  $Z_n^{\chi_j - m^{-j}\chi}$ , and results extending Theorems 2.1–2.3 without assuming (A4) follow. We leave the details to the reader.

### Acknowledgements

I thank Peter Jagers and Olle Nerman for helpful comments.

### References

- [1] ASMUSSEN, S. AND HERING, H. (1983). *Branching Processes*. Birkhäuser, Boston, MA.
- [2] BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.
- [3] BOHR, H. (1925). Zur Theorie der fastperiodischen Funktionen, I. *Acta Math.* **45**, 29–127.
- [4] DUNFORD, N. AND SCHWARTZ, J. T. (1958). *Linear Operators. I. General Theory*. Interscience, London.
- [5] DUREN, P. L. (1970). *Theory of  $H^p$  Spaces*. Academic Press, New York.
- [6] HALL, P. AND HEYDE, C. C. (1980). *Martingale Limit Theory and Its Applications*. Academic Press, New York.
- [7] JAGERS, P. (1975). *Branching Processes with Biological Applications*. John Wiley, London.
- [8] JANSON, S. (1997). *Gaussian Hilbert Spaces*. Cambridge University Press.
- [9] JANSON, S. (2004). Functional limit theorems for multitype branching processes and generalized Pólya urns. *Stoch. Process. Appl.* **110**, 177–245.
- [10] NERMAN, O. (1981). On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitsth.* **57**, 365–395.
- [11] POUYANNE, N. (2008). An algebraic approach to Pólya processes. *Ann. Inst. H. Poincaré Prob Statist.* **44**, 293–323.

SVANTE JANSON, *Uppsala University*

Department of Mathematics, Uppsala University, PO Box 480, SE-751 06 Uppsala, Sweden.